République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la

recherche scientifique

UNIVERSITE FERHAT ABBAS-SETIF

MEMOIRE

Présenté à la faculté des Sciences

Département de Mathématiques

Pour l'obtention du diplôme de

MAGISTER

Option : Mathématiques de décision

Par

Melle : GUERRA Loubna

Thème

Etude théorique et numérique de quelques méthodes de points intérieurs pour l'optimisation quadratique semi-définie

Soutenu le : 15/12/2011

Devant le jury composé :

Président :	**Mr BENTERKI Djamel**	**Prof**	**UFA Sétif**
Encadreur :	**Mr ACHACHE Mohamed**	**Prof**	**UFA Sétif**
Examinateur :	**Mr MERIKHI Bachir**	**M.C**	**UFA Sétif**
	Mme KEBBICHE Zakia	**M.C**	**UFA Sétif**

Table des matières

1 Calcul matriciel 8

1.1 Produit scalaire et normes 8

1.2 Matrices (semi-) dé…nies positives 10

1.3 Théorème de décomposition en valeurs singulières (en anglais SVD) . . . 12

1.4 Produit de Kronecker et produit de Kronecker symétrisé et ses propriétés 12

2 Programmation quadratique semi-dé…nie et méthodes de points intérieurs 16

2.1 Position du problème 16

2.2 L'importance de la programmation quadratique semi-dé…nie 17

2.3 Dualité 18

2.3.1 Formulation du problème dual 18

2.3.2 Dualité faible 20

2.3.3 Conditions d'optimalité nécessaires et su¢ santes de la solution optimale 21

2.3.4 Dualité et condition de complémentarité 22

2.4 Méthode de trajectoire centrale de type primal-dual pour (QSDP) 22

2.4.1 Description de l'algorithme générique 22

2.4.2 Les directions de Newton 26

2.4.3 La mesure de proximité 29

2.4.4 Algorithme 29

2.4.5 La convergence de l'algorithme et l'analyse de la complexité . . . 30

3 Implémentation numérique 44

3.1 Calcul de la direction 44

3.2 L'existence et l'unicité de la direction du déplacement 46

3.2.1 Condition su¢ sante 46

3.3 Résolution du système 50

3.4 Calcul de la matrice P 52
3.5 Calcul de la matice V 53

4 Tests numériques 54
4.1 Exemples 54

Introduction

Ces dernières années, le problème semi-dé…ni a fait l'objet de nombreuses études de recherche dans le domaine d'optimisation. Dans ce mémoire, on s'intéresse au problème quadratique semi-dé…ni (QSDP) qui est d'une part, une généralisation d'un problème linéaire, quadratique et semi-dé…ni (SDP) linéaire et d'autre part, il couvre beaucoup de problèmes en architecture et la génie électrique.

Le problème quadratique semi-dé…ni est un problème d'optimisation récent et qui acte sur l'ensemble des matrices symétriques semi dé…nies positives, donc l'ensemble des contraintes est un cône non polyédrique ce qui fait les méthodes classiques simpliciales ne sont pas valables. Pour cela, on utilise les méthodes de points intérieurs qui ont été initialisées pour la première fois par Karmarkar en 1984. Depuis cette date, de nombreux algorithmes ont été développés pour la programmation linéaire [12, 21], la programmation quadratique [1] et la programmation semi-dé…nie (SDP) [7, 8, 9, 14, 19, 22, 23, 24, 28] et ils ont ensuite été généralisés pour la programmation quadratique semi-dé…nie [17, 18, 26, 27]. Les méthodes de points intérieurs se distinguent en quatres catégories :

- Méthodes a¢nes.
- Méthodes projectives.
- Méthodes de réduction de potentiel.
- Méthodes de trajectoire centrale.

On note que l'é¢cacité de quelques méthodes de points intérieurs a été étudiée pour résoudre (QSDP). Nie et Yuan [17], ont développé un algorithme de la réduction de potentiel pour (QSDP) et ils ont utilisé les directions de NT. Dans [18], ils ont présenté un algorithme de points intérieurs de type prédicteur-correcteur en combinant le pas de Dikin-type et le pas de Newton et ils ont utilisé les directions de NT où ces dernières sont calculées par la méthode de gradient conjugé. Dernièrement, Toh et al [26] ont présenté un algorithme de trajectoire centrale de type prédicteur-correcteur pour des classes spéciales du problème convexe (QSDP) où les directions cherchées sont calculées à partir de l'équation de Complément de Schur. Dans [25], Toh a présenté un algorithme de points intérieurs de type prédicteur-correcteur telles que les directions cherchées sont calculées à partir de l'équation d'augmentation et il a montré l'é¢cacité numérique de ses algorithmes. Wang et Bai [27], ont utilisé les fonctions noyaux pour résoudre le problème (QSDP) par une méthode de points intérieurs primal-dual. En utilisant les fonctions noyaux, Wang et Zhu [29] ont présenté un algorithme de (TC) de type primal-dual pour résoudre le problème (QSDP) basé sur le schéma de symétrisation de NT.

Actuellement, les méthodes de trajectoire centrale de type primal-dual sont les meilleures méthodes de points intérieurs car ces algorithmes admet un bon comportement théorique telle que la complexité polynomiale et d'autre part, ils sont de type Newton ce qui conduit à une efficacité numérique. Dans [8, 9], De Klerk et al; ont proposé un algorithme de trajectoire centrale de type primal-dual à petit pas pour résoudre un problème semi-défini (SDP) basé sur la direction de Nesterov-Todd, ils ont montré que l'algorithme a une meilleure complexité polynomiale qui est de l'ordre $O(\sqrt{n}\log\frac{n}{2})$.

La résolution d'un problème quadratique semi-défini (QSDP) par la méthode de trajectoire centrale est notre objectif dans ce mémoire. Notons que dans notre cas les directions ne sont pas orthogonales (contrairement au cas (SDP)), cela conduit à une difficulté de l'analyse de la complexité. On montre aussi que notre algorithme admet une meilleure complexité polynomiale qui est de l'ordre $O(\sqrt{n}\log\frac{n}{2})$. D'autre part, cette étude est suivie par des simulations numériques intéressantes.

Le mémoire est composé de quatre chapitres :

- Dans le premier chapitre, on présente les notions fondamentales d'analyse matricielle qui sera utile par la suite du mémoire.

- Le deuxième chapitre est consacré à la définition des deux problèmes primal et dual de la programmation quadratique semi-définie et le développement d'un algorithme primal-dual à petit pas, pour tracer la trajectoire centrale approximativement. Pour determiner la direction de déplacement, on utlise le schéma de symétrisation de Nesterov-Todd et on termine ce chapitre par l'analyse de la complexité de cet algorithme.

- Dans le troisième chapitre, on démontre l'existence et l'unicité de la direction de NT ainsi son calcul.

- Dans le dernier chapitre, on présente des résultats numériques issus de l'application de cet algorithme sur des problèmes de la programmation quadratique semi-définie.

On terminera ce mémoire par une conclusion générale et perspectives.

Notations

$\mathbb{R}^n$	:	L'espace des vecteurs réels de dimension n;
$\mathbb{R}^n_+$	:	L'orthant positif de $\mathbb{R}^n$;
$\mathbb{R}^{m\times n}$	:	L'espace vectoriel des matrices réelles de taille $(m\times n)$;
$A^\top$	:	Le transposé d'une matrice $A\in\mathbb{R}^{m\times n}$;
a_{ij}	:	Elément de la matrice $A\in\mathbb{R}^{m\times n}$;
S^n	$=$	$\{X : X\in\mathbb{R}^{n\times n};\ X = X^\top\}$;
$A\succeq 0\ (A\succ 0)$	:	A est une matrice semi définie positive (définie positive),
$A\preceq 0\ (A\prec 0)$	:	A est une matrice semi définie négative (définie négative),
S^n_+	$=$	$\{X : X\in S^n;\ X\succeq 0\}$;
S^n_{++}	$=$	$\{X : X\in S^n;\ X\succ 0\}$;
$\lambda_i(A)$	:	La $i^{\text{ème}}$valeur propre de $A\in\mathbb{R}^{n\times n}$;
$sp(A)$	:	Le spectre de la matrice A;
$\lambda_{\max}(A)$	$=$	$\max\limits_i \lambda_i(A)$; si $\lambda_i(A)\in\mathbb{R};\forall i$;
$\lambda_{\min}(A)$	$=$	$\min\limits_i \lambda_i(A)$; si $\lambda_i(A)\in\mathbb{R};\forall i$;
$\mathrm{Tr}\,(A)$	$=$	$\sum\limits_i a_{ii} = \sum\limits_i \lambda_i(A)$; (la trace d'une matrice $A\in\mathbb{R}^{n\times n}$);
$\det(A)$	$=$	$\prod\limits_i \lambda_i(A)$; (determinant de $A\in\mathbb{R}^{n\times n}$);
$\rho(A)$	$=$	$\max\limits_i \lvert\lambda_i(A)\rvert$; (le rayon spectral de A),
$\lVert A\rVert_F^2$	$=$	$\mathrm{Tr}\,(AA^\top) = \sum\limits_i\sum\limits_j a_{ij}^2$;
$\lVert A\rVert_2$	$=$	$\sqrt{\rho(A^\top A)}$; la norme spectrale de A;
$A\bullet B$	$=$	$\mathrm{Tr}\,(A^\top B);\forall A,B\in\mathbb{R}^{n\times n}$;
$A^{\frac12}$	:	La racine carrée de la matrice $A\succ 0$;
A^{-1}	:	L'inverse d'une matrice régulière A;
$A\sim B$	:	Il existe une matrice inversible P telle que $A = PBP^{-1}$; (semblable),
I_n	:	La matrice identité d'ordre n;
$0_{m\times n}$	:	La matrice zéro $(m\times n)$;
$\mathrm{diag}(x)$	$=$	X la matrice diagonale avec $X_{ii} = x_i$;
$X^{(k)}$	:	Le $k^{\text{ième}}$ terme d'une suite des matrices,
e	:	Un vecteur de $\mathbb{R}^n$ tel que $e_i = 1;\forall i = 1;\ldots;n$;
$\bar{n}$	$=$	$\frac{n(n+1)}{2}$; pour tout $n\in\mathbb{N}$;
$f(x) = O(g(x))$	$\Leftrightarrow$	$\exists k>0 : f(x)\le kg(x)$;

$$\mathrm{vec}(A) = [a_{11}, a_{21}, \ldots, a_{n1}, a_{12}, a_{22}, \ldots, a_{n2}, a_{31}, a_{32}, \ldots, a_{nn}]^\top \in \mathbb{R}^{n^2}, \text{ pour } A \in \mathbb{R}^{n\times n},$$

$$(\mathrm{svec}(A))_k = \begin{cases} a_{ij} & \text{si } i = j \\ \sqrt{2}a_{ij} & \text{si } i > j \end{cases}, \text{ avec } A \in S^n \text{ et } \mathrm{svec}(A) \in \mathbb{R}^{\bar{n}}$$

$$\text{où } k = i + (j-1)n - \sum_{s=1} (s-1) \text{ pour tout } i \geq j \text{ et } i, j = 1, \ldots, n,$$

$$(\mathrm{smat}(v))_{ij} = \begin{cases} v_k & \text{si } i = j \\ \frac{1}{\sqrt{2}}v_k & \text{si } i \neq j \end{cases}, \text{ avec } v \in \mathbb{R}^{\bar{n}} \text{ et } \mathrm{smat}(v) \in S^n$$

$$\text{où } k = i + (j-1)n - \sum_{s=1} (s-1) \text{ pour tout } i \geq j \text{ et } i, j = 1, \ldots, n.$$

Terminologie

(QSDP)	:	Programmation quadratique semi-définie,
(SDP)	:	Programmation linéaire semi-définie,
(PQSDP)	:	Le problème quadratique semi-défini sous forme primal,
(DQSDP)	:	Le problème dual de (PQSDP),
(TC)	:	La trajectoire centrale,
CPI	:	Condition de points intérieurs,
K.K.T	:	Karush-Kuhn-Tucker,
NT	:	Nesterov et Todd,
SVD	:	La décomposition en valeurs singulières,
c-à-d	:	C'est à dire,
$\nabla f(X)$	:	Le gradient de f au point X, où $f : S^n \to \mathbb{R}$,
$\nabla^2 f(X)$	:	Le Hessien de f au point X, où $f : S^n \to \mathbb{R}$,
$\frac{\partial f(X)}{\partial x_{ij}}$	:	Les dérivées partielles de f au point x_{ij}, où $X \in S^n$,
C^k	:	Espace des fonctions k fois continûment différentiables.

Chapitre 1

Calcul matriciel

1.1 Produit scalaire et normes

Définition 1.1.1 *Le produit scalaire usuel de deux vecteurs* x *et* y *de* $\mathbb{R}^n$ *est défini par :*

$$\langle x, y\rangle = \sum_{i=1} x_i y_i = x^\top y.$$

De même, on définit un produit scalaire sur l'ensemble des matrices carrées réelles :

Définition 1.1.2 *Soient* $A, B \in \mathbb{R}^{n\times n}$*, le produit scalaire de* A *et* B *noté* $A \bullet B$ *est défini par :*

$$A \bullet B = Tr\,(A^\top B) = \sum_{i=1}\sum_{j=1} a_{ij} b_{ij} = B \bullet A,$$

où $Tr\,(.)$ *vérifie les propriétés suivantes :*

Propriétés de la trace

1. $\forall A, B \in \mathbb{R}^{n\times n}; \forall \alpha \in \mathbb{R} : Tr\,(A + \alpha B) = Tr\,(A) + \alpha Tr\,(B)$; (linéarité),
2. $\forall A, B \in \mathbb{R}^{n\times n} : Tr\,(AB) = Tr\,(BA)$;
3. $\forall A \in \mathbb{R}^{n\times n} : Tr\,(A) = Tr\,(A^\top)$;
4. $\forall A, B \in \mathbb{R}^{n\times n} : A \sim B \Rightarrow Tr\,(A) = Tr\,(B)$.

Enonçons maintenant les propriétés de produit scalaire $A \bullet B$. Pour cela, on considère $A, B, C \in \mathbb{R}^{n\times n}$ et $\lambda, \alpha \in \mathbb{R}$ alors :

1. $\langle A + B, C\rangle = \langle A, C\rangle + \langle B, C\rangle;$

2. $\langle \lambda A, \alpha B\rangle = \lambda\alpha\langle A, B\rangle;$

3. $\langle A, B\rangle = \langle B, A\rangle.$

Définition 1.1.3 *La norme vectorielle est une application de $\mathbb{R}^n$ dans $\mathbb{R}_+$, notée par $\|.\|$ et vérifie les conditions suivantes :*

1. $\forall x \in \mathbb{R}^n : \|x\| = 0 \iff x = 0;$

2. $\forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n : \|\lambda x\| = |\lambda| \|x\|;$

3. $\forall x, y \in \mathbb{R}^n : \|x + y\| \leq \|x\| + \|y\|.$

Définition 1.1.4 *Soient $A, B \in \mathbb{R}^{n\times n}$, l'application $\|.\| : \mathbb{R}^{n\times n} \to \mathbb{R}_+$ est appelée norme matricielle si elle vérifie les conditions suivantes :*

1. $\|A\| = 0 \iff A = 0;$

2. $\|\alpha A\| = |\alpha| \|A\|, \alpha \in \mathbb{R};$

3. $\|A + B\| \leq \|A\| + \|B\|;$

4. $\|AB\| \leq \|A\|\|B\|.$

Définition 1.1.5 *Soit $\|.\|_\alpha$ une norme vectorielle sur $\mathbb{R}^n$, on appelle une norme matricielle subordonnée associée à la norme vectorielle $\|.\|_\alpha$, toute application de $\mathbb{R}^{n\times n}$ dans $\mathbb{R}_+$ qui associe à chaque matrice A, le nombre $\|A\|$ qui est défini par :*

$$\|A\| = \sup_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} \frac{\|Ax\|_\alpha}{\|x\|_\alpha}.$$

Remarque 1.1.1 *Pour toute norme matricielle subordonnée on a :*

1. $\|Ax\|_\alpha \leq \|A\|\|x\|_\alpha, \forall x \neq 0, x \in \mathbb{R}^n.$

2. $\exists x \neq 0, x \in \mathbb{R}^n, (\exists x : \|x\|_\alpha = 1) : \|Ax\|_\alpha = \|A\|.$

3. La norme matricielle subordonnée est une norme matricielle sur $\mathbb{R}^{n\times n}$.

Lemme 1.1.1 *Pour toute matrice $A \in \mathbb{R}^{n\times n}$, on a :*

1. $\|A\|_1 = \max_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}| \right);$

2. $\|A\|_\infty = \max_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}| \right);$

3. $\|A\|_2 = \sqrt{\rho(A^\top A)};$ est appelée la norme spectrale.

On utilisera également la norme de Frobenius :

$$\|A\|_F = \sqrt{A \bullet A} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij})^2}.$$

On note que si : $A = A^\top$; alors on obtient facilement les résultats suivants :

$$\|A\|_F = \sqrt{Tr\,(A^\top A)} = \sqrt{Tr\,(A^2)} = \sqrt{\sum_{i=1}^{n} \lambda_i^2(A)};$$

et

$$\|A\|_2 = \sqrt{\rho(A^2)} = \max_{i=1}^{n} |\lambda_i(A)|;$$

de plus,

$$\|A\|_2 \le \|A\|_F \le \sqrt{n}\,\|A\|_2;$$

et pour toute norme matricielle on a :

$$\rho(A) \le \|A\|.$$

1.2 Matrices (semi-) définies positives

Définition 1.2.1 *Soit* $A \in \mathbb{R}^{n \times n}$.
A *est dite semi-définie positive si et seulement si :* $x^\top A x \ge 0, \forall x \in \mathbb{R}^n.$
A *est dite définie positive si et seulement si :* $x^\top A x > 0, \forall x \in \mathbb{R}^n, x \neq 0.$

Théorème 1.2.1 *Soit* $A \in S^n$*, les conditions suivantes sont équivalentes :*

1. $A \in S^n_+$ (resp $A \in S^n_{++}$),

2. $\lambda_{\min}(A) \geq 0$ ($\lambda_{\min}(A) > 0$),

3. $\exists P \in \mathbb{R}^{n \times n} : A = P^{\top}P$ (resp $\exists P \in \mathbb{R}^{n \times n} : rg(P) = n; A = P^{\top}P$).

Proposition 1.2.1 : *Soit* $A \in S^n_{++}$; *alors il existe une matrice unique* $B \in S^n_{++}$ *telle que* $A = B^2$; *et on la note souvent par* $B = A^{\frac{1}{2}}$. *De plus,* B *est appelée la racine carrée de* A.

Dans la suite de cette partie, on s'intéresse aux matrices symétriques semi-définies positives. On définit sur S^n l'ordre de Löwner par :

$$A \succeq B \Leftrightarrow A - B \succeq 0; (\text{resp } A \succ B \Leftrightarrow A - B \succ 0),$$

et on a la proposition suivante :

Proposition 1.2.2 : *Soient* $A; B \in S^n_+$, *alors :*

1. $A + B \succeq B$.

2. $A^{\frac{1}{2}} B A^{\frac{1}{2}} \succeq 0$.

3. $Tr(AB) \leq Tr(A) Tr(B)$.

4. $Tr(AB) \geq 0$.

Les prochains lemmes seront utiles par la suite du mémoire.

Lemme 1.2.1 : *Soient* $A; B \in S^n_+$; *les assertions suivantes sont équivalentes :*

1. $A \bullet B = 0$;

2. $AB = 0$;

3. $\frac{1}{2}(AB + BA) = 0$.

Lemme 1.2.2 : *Soient* $A; B \in S^n_+$. *Alors :*

$$\lambda_{\min}(A)\lambda_{\max}(B) \leq \lambda_{\min}(A) Tr(B) \leq A \bullet B \leq \lambda_{\max}(A) Tr(B) \leq n\lambda_{\max}(A)\lambda_{\max}(B).$$

1.3 Théorème de décomposition en valeurs singulières (en anglais SVD)

La décomposition en valeurs singulières (SVD) permet la généralisation de la notion de valeurs propres aux matrices rectangulaires de taille $(m \times n)$ et la construction d'inverses généralisés ([6], [20]).

Lemme 1.3.1 *Soit* $A \in \mathbb{R}^{m \times n}$; *alors la matrice* $A^\top A \in S^n_+$ *.*

Définition 1.3.1 *Soit* $A \in \mathbb{R}^{m \times n}$. *Les valeurs singulières d'une matrice* A *sont les racines carrées des valeurs propres de* $A^\top A$.

Théorème 1.3.1 *Soit* $A \in \mathbb{R}^{m \times n}$ *avec les valeurs propres de* $A^\top A$ *non toutes nulles. Il existe deux matrices unitaires* $U \in \mathbb{R}^{m \times m}$ *et* $V \in \mathbb{R}^{n \times n}$ *et une matrice* $\Sigma \in \mathbb{R}^{m \times n}$ *telles que :*

1. $U^\top A V = \Sigma$.
2. $\Sigma = \begin{bmatrix} D & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$ où $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ avec $r = \min(m, n)$ et $\sigma_1 \geq \sigma_2 > \ldots > \sigma_r > 0$ sont les valeurs singulières non nulles de A.
3. Les vecteurs colonnes $v_1, \ldots, v_r$ de V sont des vecteurs propres de $A^\top A$ associés à $\sigma_1^2, \ldots, \sigma_r^2$.

   Les vecteurs colonnes $u_1, \ldots, u_r$ de U sont des vecteurs propres de $AA^\top$ associés à $\sigma_1^2, \ldots, \sigma_r^2$.

Nous introduisons maintenant le produit de Kronecker symétrisé et ses propriétés (voir [15]) :

1.4 Produit de Kronecker et produit de Kronecker symétrisé et ses propriétés

Définition 1.4.1 *Soient* $A \in \mathbb{R}^{m \times n}$ *et* $B \in \mathbb{R}^{k \times l}$, *le produit de Kronecker de* A *et* B, *noté* $A \otimes B \in \mathbb{R}^{mk \times nl}$ *est défini par :*

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \ldots & a_{mn}B \end{bmatrix}.$$

Proposition 1.4.1 *Soient* A, B, C, D *quatre matrices avec des tailles adaptées. Alors :*

1. $(A \otimes B)^\top = A^\top \otimes B^\top$.
2. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.
3. $\operatorname{vec}(ABC) = \left(C^\top \otimes A\right) \operatorname{vec}(B)$.
4. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
5. Si A et B sont symétriques et définies positives, alors $(A \otimes B)$ l'est aussi.

Définition 1.4.2 *Soient* $A, B \in \mathbb{R}^{n \times n}$ *et* $X \in S^n$*, le produit de Kronecker symétrisé de* B *et* A*, noté par* $(B \otimes_s A)$ *est une matrice de* $\mathbb{R}^{\tilde{n} \times \tilde{n}}$ *qui vérifie*

$$(B \otimes_s A)\operatorname{svec}(X) = \operatorname{svec}\left(\frac{1}{2}(AXB^\top + BXA^\top)\right), \forall X \in S^n,$$

où

$$\operatorname{svec}(X) = \left[X_{11}, \sqrt{2}X_{12}, \ldots, \sqrt{2}X_{1n}, X_{22}, \sqrt{2}X_{23}, \ldots, X_{nn}\right]^\top, \text{ pour } X \in S^n \text{ et } \tilde{n} = \frac{n(n+1)}{2}.$$

On peut aussi calculer $(B \otimes_s A)$ *par l'expression suivante :*

$$(B \otimes_s A) = \frac{1}{2}U^\top(B \otimes A + A \otimes B)U,$$

avec $U \in \mathbb{R}^{n^2 \times \tilde{n}}$ *et vérifie :*

$$U\operatorname{svec}(X) = \operatorname{vec}(X), \forall X \in S^n.$$

Pour calculer U*, on considère* $\hat{e}_i$ *le* $i^{\text{ème}}$ *vecteur de* $\mathbb{R}^{\tilde{n}}$*, pour tout* $i = 1, \ldots, \tilde{n}$*, et*

$$E_i = \operatorname{smat}(\hat{e}_i) \in S^n.$$

Alors E_i *est de la forme :*

$$E_i = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & \frac{1}{\sqrt{2}} & \cdots & 0 \\ \vdots & & & \ddots & \vdots & & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{2}} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \quad \textit{ou} \quad E_i = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix};$$

et la $i^{\text{ème}}$*colonne de* U *est*

$$\hat{u}_i = U\hat{e}_i = U\mathrm{svec}(E_i) = \mathrm{vec}(E_i) = \mathrm{vec}(\mathrm{smat}(\hat{e}_i));$$

alors

$$U = \mathrm{vec}\left[\mathrm{smat}(\hat{e}_1); \ldots; \mathrm{smat}(\hat{e}_{\hat{n}})\right].$$

Par exemple, si $n = 2$ *alors* $\hat{n} = 3$ *et les vecteurs* $\hat{e}_i;\ i = 1;2;3$ *sont donnés par :*

$$\hat{e}_1 = [1;\,0;\,0]^\top;\ \hat{e}_2 = [0;\,1;\,0]^\top;\ \hat{e}_3 = [0;\,0;\,1]^\top;$$

donc

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};\ E_2 = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{bmatrix};\ E_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix};$$

et par conséquent

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 3}.$$

Proposition 1.4.2 *Soient* $A; B; C; D$ *quatre matrices avec des tailles adaptées et* $X \in S^n$*. Alors :*

1. $(B \otimes_s A)\mathrm{svec}(X) = \mathrm{svec}(\frac{1}{2}(AXB^\top + BXA^\top)).$
2. $(B \otimes_s A)^\top = B^\top \otimes_s A^\top.$

3. $B \otimes_s A = A \otimes_s B$:

4. $(B \otimes_s A)(D \otimes_s C) = \frac{1}{2}(BD \otimes_s AC + BC \otimes_s AD)$:

5. Si A et B sont symétriques et dé…nies positives, alors $(A \otimes_s B)$ l'est aussi.

Chapitre 2

Programmation quadratique semi-définie et méthodes de points intérieurs

2.1 Position du problème

Soient deux entiers naturels m, n $(m \geq 1)$, un vecteur $b \in \mathbb{R}^m$, des matrices $C, A_i \in S^n$, $i = 1, \ldots, m$ et $Q : S^n \to S^n$ un opérateur linéaire auto-adjoint c'est à dire :

$$\forall A, B \in S^n : A \bullet Q(B) = Q(A) \bullet B.$$

Un programme quadratique semi-défini sous forme primal est un problème d'optimisation donné par :

$$(PQSDP) \quad \begin{cases} \min_X & C \bullet X + \frac{1}{2} X \bullet Q(X) \\ \text{sujet à :} & A_i \bullet X = b_i, \; i = 1, \ldots, m; \quad X \in S^n_+. \end{cases}$$

On note par la suite

$$\begin{aligned} F_P &= \{X \in S^n_+ : A_i \bullet X = b_i, \; i = 1, \ldots, m\}, \\ F_P^+ &= \{X \in S^n_{++} : A_i \bullet X = b_i, \; i = 1, \ldots, m\}, \end{aligned}$$

l'ensemble des solutions réalisables (et strictement réalisables, respectivement) pour (PQSDP).

La valeur optimale primale du problème $(PQSDP)$ est définie par :

$$m_P = \inf_X \{C \bullet X + \frac{1}{2} X \bullet Q(X) : A_i \bullet X = b_i;\ i = 1; \dots ; m;\ X \in S^n_+\}.$$

On dit que X^* est une solution optimale primale de $(PQSDP)$ si :

$$X^* \in F_P \text{ et } m_P = C \bullet X^* + \frac{1}{2} X^* \bullet Q(X^*).$$

2.2 L'importance de la programmation quadratique semi-définie

La programmation quadratique semi-définie (QSDP) est l'un des problèmes d'optimisation les plus étudiés ces dernières années.

L'intérêt pour ce type de problème est justifié par les facteurs suivants :

• Le problème (QSDP) contient des classes spéciales et importantes comme la programmation linéaire semi-définie (SDP), la programmation linéaire et la programmation quadratique.

• L'existence de plusieurs applications du problème (QSDP) en différents domaines comme : l'architecture et la génie électrique...

On note que si $Q(X) = 0_{n \times n}$ alors (QSDP) devient un problème de programmation semi-définie linéaire (SDP).

De plus, plusieurs problèmes d'optimisation convexe peuvent être transformer au problème (QSDP) (pour plus de détails voir [10], [25], [26]). En particulier, le problème "The nearest correlation matrix (NCM) problem" est un exemple de (QSDP) qui est défini par :

$$\min_X \{\frac{1}{2} \|L(X - K)\|_F^2 : \operatorname{diag}(X) = e;\ X \in S^n_+\}; \tag{2.1}$$

où $K \in S^n$ une matrice donnée et L un opérateur linéaire auto-adjoint ($L : S^n \to S^n$). D'une part, le problème (2.1) est un cas particulier d'un problème des moindres carrés semi-défini (SDLS) qui est donné par :

$$\min_X \{\left\| L(X) - \hat{K} \right\|_F^2 : \operatorname{diag}(X) = b;\ X \in S^n_+\};$$

où $L : S^n \to S^p$ un opérateur linéaire auto-adjoint et $\hat{K} \in S^p$ une matrice donnée, on prend souvent $\hat{K} = L(K)$ avec $K \in S^n$ une matrice donnée. Afin d'obtenir le problème (NCM), on prend $p = n$, $L(X) = U^{\frac{1}{2}} X U^{\frac{1}{2}}$ avec $U \in S^n_{++}$ et $b = e$.

D'autre part, le problème (QSDP) résulte du (2.1) avec $C = -L^2(K)$ et pour $i = 1, \ldots, n$, la matrice A_i est définie par :

$$A_i[j,k] = \begin{cases} 1 & \text{si } i = j = k, \\ 0 & \text{ailleurs,} \end{cases}$$

et $b = e$. Plusieurs recherches en (QSDP) basées sur le problème de (NCM) avec un choix spécial : $L = U^{\frac{1}{2}} \circledast U^{\frac{1}{2}}$ (ainsi $Q = U \circledast U$) pour une matrice donnée $U \in S_+^n$ où l'opérateur $\circledast$ est défini par :

$$\begin{array}{rrcl} A \circledast B : & \mathbb{R}^{n \times l} & \to & S^q \\ & M & \mapsto & A \circledast B(M) = (BMA^\top + AM^\top B^\top)/2, \end{array}$$

avec $A \in \mathbb{R}^{q \times l}$ et $B \in \mathbb{R}^{q \times n}$.

L'un des premiers travaux qui traite (2.1), a été fait par Higham [10], qui a proposé "Modified alternating projection solution method" et il a considéré le problème (2.1) avec $L(X) = \Sigma \circ X$ pour une matrice donnée $\Sigma \in S^n \cap \mathbb{R}_+^{n \times n}$ (respectivement $Q(X) = U \circ X$ avec $U = \Sigma \circ \Sigma$) où $\Sigma \circ X$ désigne le produit de Hadamard qui est défini par : $(\Sigma \circ X)_{ij} = \Sigma_{ij} X_{ij}$, avec Σ et X deux matrices de même taille.

2.3 Dualité

2.3.1 Formulation du problème dual

Soit le problème d'optimisation (PQSDP) primal standard :

$$\begin{array}{ll} \min_X & C \bullet X + \frac{1}{2} X \bullet Q(X) \\ \text{sujet à :} & A_i \bullet X = b_i,\ i = 1, \ldots, m, \quad X \in S_+^n. \end{array}$$

Pour obtenir le problème dual de (PQSDP), on considère la fonction Lagrangienne $L : S^n \times \mathbb{R}^m \to \mathbb{R}$ définie par :

$$L(X, y) = C \bullet X + \frac{1}{2} X \bullet Q(X) + \sum_{i=1}^{m} (b_i - A_i \bullet X) y_i,$$

le dual de (PQSDP) est donné par :

$$\{\max_{y \in \mathbb{R}^m} \min_{X \in S_+^n} L(X, y) = \max_{y \in \mathbb{R}^m} q(y),$$

où

$$
\begin{aligned}
q(y) &= \min_{X \in S^n_+} L(X;y),\; y \in \Re^m; \\
&= \min_{X \in S^n_+} \left\{ C \bullet X + \frac{1}{2} X \bullet Q(X) + \sum_{i=1}^{m} (b_i - A_i \bullet X) y_i \right\} ; y \in \Re^m; \\
&= \min_{X \in S^n_+} \left\{ \left(C - \sum_{i=1}^{m} y_i A_i + Q(X) \right) \bullet X + b^{\top} y - \frac{1}{2} X \bullet Q(X) \right\} ; y \in \Re^m.
\end{aligned}
$$

Finalement, on obtient :

$$
q(y) = \begin{cases} b^{\top} y - \frac{1}{2} X \bullet Q(X) & \text{si } \left(C - \sum_{i=1}^{m} y_i A_i + Q(X) \right) \succeq 0 \\ -\infty & \text{sinon} \end{cases} .
$$

Donc le dual est un programme quadratique semi-défini donné par :

$$
(DQSDP) \quad \begin{cases} \max_{(X;y;Z)} b^{\top} y - \frac{1}{2} X \bullet Q(X) \\ \sum_{i=1}^{m} y_i A_i - Q(X) + Z = C \\ X; Z \in S^n_+ \text{ et } y \in \Re^m \end{cases} ;
$$

où l'inconnu est $(X; y; Z) \in S^n_+ \times \Re^m \times S^n_+$.
On note par la suite

$$
\begin{aligned}
F_D &= \{X; Z \in S^n_+ \text{ et } y \in \Re^m : \sum_{i=1}^{m} y_i A_i - Q(X) + Z = C\}; \\
F_D^+ &= \{X; Z \in S^n_{++} \text{ et } y \in \Re^m : \sum_{i=1}^{m} y_i A_i - Q(X) + Z = C\};
\end{aligned}
$$

l'ensemble des solutions réalisables (et strictement réalisables, respectivement) pour $(DQSDP)$. La valeur optimale duale du problème $(DQSDP)$ est définie par :

$$
m_D = \sup_{(X;y;Z)} \left\{ b^{\top} y - \frac{1}{2} X \bullet Q(X) : \sum_{i=1}^{m} y_i A_i + Z - Q(X) = C;\; y \in \Re^m \text{ et } X; Z \in S^n_+ \right\} .
$$

On dit que $(X^*; y^*; Z^*)$ est une solution optimale duale de $(DQSDP)$ si :

$$
(X^*; y^*; Z^*) \in F_D \text{ et } m_D = b^{\top} y^* - \frac{1}{2} X^* \bullet Q(X^*).
$$

2.3.2 Dualité faible

On a le résultat suivant :

Proposition 2.3.1 : *Pour toute* $X \in F_P$ *et* $(X; y; Z) \in F_D$ *on a :*

$$Z \bullet X = C \bullet X - b^\top y + X \bullet Q(X) \geq 0.$$

Preuve : Soient $X \in F_P$ et $(X; y; Z) \in F_D$; alors :

$$\begin{aligned}
C \bullet X + X \bullet Q(X) - b^\top y &= (C + Q(X)) \bullet X - \sum_{i=1}^{m} b_i y_i, \\
&= (C + Q(X)) \bullet X - \sum_{i=1}^{m} y_i \langle A_i, X \rangle, \\
&= \left(C + Q(X) - \sum_{i=1}^{m} y_i A_i \right) \bullet X, \\
&= Z \bullet X \geq 0, \text{ car } X \text{ et } Z \in S^n_+,
\end{aligned}$$

la dernière inégalité découle de la Proposition [1.2.2]. Ce qui achève la preuve. ■

Remarque 2.3.1 *Soient* $X \in F_P$ *et* $(X; y; Z) \in F_D$; *alors la différence*

$$C \bullet X + X \bullet Q(X) - b^\top y$$

est appelée le saut de dualité.

Pour résoudre les deux problèmes (PQSDP) et (DQSDP), on suppose qu'ils vérifient les hypothèses suivantes :

Hypothèse 1. Condition d'indépendance : Les matrices $[A_1; A_2; \ldots; A_m]$ sont linéairement indépendantes.

Hypothèse 2. Condition de points intérieurs : (CPI) On suppose qu'il existe un point primal-dual strictement réalisable $(X^0; y^0; Z^0)$, en autre terme ce point vérifie

$$\begin{cases}
A_i \bullet X^0 = b_i, \quad i = 1, \ldots, m \\
\sum_{i=1}^{m} y_i^0 A_i - Q(X^0) + Z^0 = C, \\
X^0, Z^0 \in S^n_{++} \text{ et } y^0 \in \mathbb{R}^m.
\end{cases}$$

Hypothèse 3. La monotonie. L'opérateur linéaire Q est monotone c-à-d :

$$X \bullet Q(X) \geq 0, \forall X \in S^n.$$

2.3.3 Conditions d'optimalité nécessaires et suffisantes de la solution optimale

Pour étudier la convexité et la différentiabilité de (QSDP), on a besoin du lemme suivant :

Lemme 2.3.1 *Soit* Q *un opérateur linéaire auto-adjoint et monotone, alors :* $\nabla Q(X)$ *est aussi un opérateur monotone* $\forall X \in S^n$.

Preuve : Soient $X \in S^n$ et Q un opérateur linéaire auto-adjoint et monotone, alors Q est différentiable et sa dérivée est lui même c-à-d

$$\nabla Q(X)Y = Q(Y), \forall Y \in S^n;$$

car : si on prend $Y \in S^n$ telle que $|\lambda_i(Y)|$ est suffisamment petite (voir [28]) alors :

$$Q(X+Y) \simeq Q(X) + \nabla Q(X)(Y), \forall X, Y \in S^n;$$

qui est équivalent à

$$Q(X+Y) - Q(X) = Q(Y) \simeq \nabla Q(X)(Y), \forall X, Y \in S^n;$$

donc

$$\nabla Q(X)Y \bullet Y = Q(Y) \bullet Y \geq 0, \forall Y \in S^n \text{ car } Q \text{ est monotone,}$$

et par conséquent $\nabla Q(X)$ est un opérateur monotone $\forall X \in S^n$. Ce qui achève la preuve. ■

On rappelle que le problème (PQSDP) consiste à minimiser la fonction :

$$f(X) = C \bullet X + \frac{1}{2} X \bullet Q(X);$$

sous l'ensemble des contraintes

$$F_P = \{A_i \bullet X = b_i, i = 1, \dots, m, X \in S^n_+\};$$

qui s'écrit comme l'intersection d'un cône convexe non polyédrique noté par ($X \in S^n_+$) et des contraintes linéaires $\{A_i \bullet X = b_i, i = 1, \dots, m\}$. Donc F_P est convexe et d'après la condition de points intérieurs (H_2) on déduit que l'ensemble F_P est d'intérieur relatif non vide, alors la condition de Slater est satisfaite (c-à-d F_P est convexe et $F_P^+ \neq \emptyset$). De plus, on peut montrer que la fonction objective $f(X)$ est deux fois différentiable et

convexe c-à-d $\nabla^2 f(X)H \bullet H = \nabla Q(X)H \bullet H \geq 0$, $\forall X, H \in S^n$, où la dernière inégalité est déduite du Lemme [2.3.1].
Par conséquent, le problème d'optimisation $(PQSDP)$ est convexe et différentiable, à contraintes qualifiées. Alors les conditions d'optimalités de $(K.K.T)$ sont nécessaires et suffisantes et s'écrivent comme suit :

$$(K.K.T) \quad \begin{cases} A_i \bullet X & = b_i;\ i = 1, \ldots, m; \\ \sum\limits_{i=1}^{m} y_i A_i + Z - Q(X) & = C; \\ Z \bullet X & = 0;\ X, Z \in S_+^n. \end{cases}$$

2.3.4 Dualité et condition de complémentarité

Le saut de dualité de problèmes $(PQSDP)$ et $(DQSDP)$ aux solutions réalisables (X, y, Z) est donné par :

$$C \bullet X + \frac{1}{2} X \bullet Q(X) - b^\top y + \frac{1}{2} X \bullet Q(X) = C \bullet X + X \bullet Q(X) - b^\top y = X \bullet Z.$$

On peut exprimer la condition pour X^* et (X^*, y^*, Z^*) d'être des solutions optimales de $(PQSDP)$ et $(DQSDP)$, sous forme d'une condition dite de complémentarité. Cela est donné dans le cadre du théorème suivant :

Théorème 2.3.1 *[16, Théorème 2.3.8] Soient* $X^* \in F_P$ *et* $(X^*, y^*, Z^*) \in F_D$*. Alors,* X^* *et* (X^*, y^*, Z^*) *sont des solutions optimales pour* $(PQSDP)$ *et* $(DQSDP)$*, si et seulement si* $X^* Z^* = 0$.

2.4 Méthode de trajectoire centrale de type primal-dual pour (QSDP)

2.4.1 Description de l'algorithme générique

Rappelons-nous, l'énoncé des problèmes primal et dual :

$$(PQSDP) \quad \begin{cases} \min C \bullet X + \frac{1}{2} X \bullet Q(X); \\ A_i \bullet X = b_i;\ i = 1, \ldots, m; \\ X \in S_+^n; \end{cases}$$

et

$$(DQSDP)\quad \begin{cases} \max b^{T}y - \frac{1}{2}X \bullet Q(X); \\ \sum_{i=1}^{m} y_i A_i + Z - Q(X) = C; \\ X; Z \in S_{+}^{n};\ y \in \mathbb{R}^{m}. \end{cases}$$

Sous les hypothèses $H_1; H_2; H_3$, il est connu que ([25],[7]) trouver une solution optimale de $(PQSDP)$ et $(DQSDP)$ est équivalent à trouver une solution du système non linéaire suivant :

$$(S)\quad \begin{cases} A_i \bullet X & = b_i;\ i = 1; \ldots ; m; \\ \sum_{i=1}^{m} y_i A_i + Z - Q(X) & = C; \\ XZ & = 0;\ X; Z \in S_{+}^{n};\ y \in \mathbb{R}^{m}. \end{cases}$$

Où la dernière équation représente la condition de complémentarité.
Afin d'introduire une méthode de points intérieurs au système (S), on associe au problème $(PQSDP)$ le problème de minimisation barrière non linéaire suivant :

$$(P_{\mu})\quad \begin{cases} \min C \bullet X + \frac{1}{2}X \bullet Q(X) - \mu \log\det X; \\ A_i \bullet X = b_i;\ i = 1; \ldots ; m; \\ X \in S_{++}^{n};\ \mu > 0; \end{cases}$$

avec la fonction objective est

$$f_{\mu}(X) = C \bullet X + \frac{1}{2}X \bullet Q(X) - \mu \log\det X.$$

Les deux lemmes suivants permettent de calculer le gradient de la fonction barrière logarithmique primale $f_{\mu}(X)$. Pour la démonstration voir l'Annexe.

Lemme 2.4.1 *[7] Soit* $f : S_{++}^{n} \to \mathbb{R}$ *une fonction donnée par* $f(X) = \log\det X$*, et*

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \cdots & \frac{\partial f(X)}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X)}{\partial x_{n1}} & \cdots & \frac{\partial f(X)}{\partial x_{nn}} \end{bmatrix};$$

alors $\nabla f(X) = X^{-1}$.

Lemme 2.4.2 *[7] Soit* $f : S_{++}^{n} \longrightarrow \mathbb{R}$ *une fonction donnée par* $f(X) = Tr(CX) = C \bullet X;$ *où* $C \in S^{n};$ *alors* $\nabla f(X) = C.$

Le résultat suivant conduit à obtenir le Hessien de la fonction $f_\mu(X)$. Pour la démonstration voir l'Annexe.

Lemme 2.4.3 *[7] Soient* $f : S^n_{++} \to \mathbb{R}$ *une fonction donnée par* $f(X) = \log\det X$ *et* $\nabla f(X) = X^{-1}$*, alors* $\nabla^2 f(X)$ *est un opérateur linéaire qui vérifie*

$$\nabla^2 f(X)H = -X^{-1}HX^{-1}; \forall H \in S^n.$$

Le résultat suivant montre la stricte convexité de la fonction barrière logarithmique primale.

Proposition 2.4.1 *Soient* $\mu > 0$ *et* $f_\mu : S^n_{++} \to \mathbb{R}$ *une fonction définie par :*

$$f_\mu(X) = C \bullet X + \frac{1}{2} X \bullet Q(X) - \mu \log\det X.$$

Alors, la fonction $f_\mu(X)$ *est strictement convexe pour tout* $\mu > 0$ *(fixé).*

Preuve : Soient $X \in S^n_{++}$; $H \in S^n$. D'après les Lemmes précédents on a :

$$\nabla f_\mu(X) = C + Q(X) - \mu X^{-1};$$

et

$$\begin{aligned} \nabla^2 f_\mu(X)H &= \nabla Q(X)H + \mu X^{-1}HX^{-1}; \\ &= Q(H) + \mu X^{-1}HX^{-1}. \end{aligned}$$

Par conséquent

$$\nabla^2 f_\mu(X)H \bullet H = Q(H) \bullet H + \mu X^{-1}HX^{-1} \bullet H;$$

puisque Q est monotone et $\mu > 0$; il suffit de montrer que :

$$X^{-1}HX^{-1} \bullet H > 0; \forall X \in S^n_{++} \text{ et } H \in S^n.$$

On a :

$$\begin{aligned} X^{-1}HX^{-1} \bullet H &= \mathrm{Tr}(X^{-1}HX^{-1}H); \\ &= \mathrm{Tr}((X^{-\frac{1}{2}}HX^{-\frac{1}{2}})^2); \\ &= \left\| X^{-\frac{1}{2}}HX^{-\frac{1}{2}} \right\|_F^2 > 0; \forall H \neq 0 \text{ et } H \in S^n. \end{aligned}$$

Donc $f_\mu(X)$ est strictement convexe. Ce qui achève la preuve. ■

Notons que la démonstration de ce Lemme est faite en [17] avec un choix spécial de l'opérateur $Q(X)$:

Enonçons maintenant les propriétés suivantes:

- $-\log\det X$ est appelée la fonction barrière logarithmique.
- Le paramètre $\mu > 0$ est appelé le paramètre de barrière.
- (P_μ) est un problème d'optimisation convexe et différentiable.
- Les conditions d'optimalité nécessaires d'ordre 1 pour que $X(\mu)$ soit une solution pour (P_μ), est l'existence d'un vecteur $y(\mu)$ tel que:

$$\begin{cases} \nabla f_\mu(X) - \sum_{i=1}^{m} y_i(\mu)A_i & = 0; \\ A_i \bullet X(\mu) & = b_i;\ i = 1;\ldots;m; \end{cases}$$

avec

$$\nabla f_\mu(X) = C + Q(X) - \mu X^{-1}.$$

On pose

$$Z(\mu) = \mu X^{-1};$$

alors

$$X(\mu)Z(\mu) = \mu I.$$

Donc les conditions nécessaires et suffisantes de (K.K.T) s'écrivent:

$$\begin{cases} \sum_{i=1}^{m} y_i(\mu)A_i + Z(\mu) - Q(X(\mu)) & = C; \\ A_i \bullet X(\mu) & = b_i;\ i = 1;\ldots;m; \\ X(\mu)Z(\mu) & = \mu I. \end{cases} \tag{2.2}$$

- L'existence et l'unicité de la solution $(X(\mu); y(\mu); Z(\mu))$ pour tout $\mu > 0$ (μ fixé) reviennent à l'existence et l'unicité de la solution du problème (P_μ) (voir [17]).
- L'ensemble $\mathcal{C} = \{(X(\mu); y(\mu); Z(\mu)); \mu > 0\}$ est appelé la trajectoire centrale, si μ tend vers zéro alors $(X(\mu); y(\mu); Z(\mu))$ devient une solution du système (S) c-à-d des solutions optimales de deux problèmes (PQSDP) et (DQSDP), (voir [7], [17]).
- La courbe $\mathcal{C}$ est lisse pour tout $\mu > 0$.
- Le système d'équations non linéaires (2.2) caractérise la solution optimale du pro-

blème (P_1), donc la résolution de ce problème est équivalente à la résolution de (2:2), et pour résoudre ce dernier, on utilise la méthode de trajectoire centrale de type primal-dual.

Stratégie de l'algorithme

- On commence par $(X^0; y^0; Z^0)$ un point strictement réalisable primal-dual et apartenant à un voisinage de trajectoire centrale, avec μ^0 connu $(\mu^0 = \frac{X^0 \bullet Z^0}{n})$:
- On construit le nouveau triple

$$\begin{cases} X^{(k+1)} &= X^{(k)} + \Delta X^{(k)}; \\ y^{(k+1)} &= y^{(k)} + \Delta y^{(k)}; \\ Z^{(k+1)} &= Z^{(k)} + \Delta Z^{(k)}; \end{cases}$$

où $(\Delta X^{(k)}; \Delta y^{(k)}; \Delta Z^{(k)})$ sont les directions de Newton.
- On vérifie que $X^{(k+1)}; Z^{(k+1)}$ sont strictement réalisables pour tout $\mu > 0$; et on distingue deux cas :
Si $n\mu^{(k)} < \varepsilon$ où ε est la précision donnée, alors $X^{(k+1)}; y^{(k+1)}$ et $Z^{(k+1)}$ sont des solutions approchées du système (2.2).
Sinon $(n\mu^{(k)} \geq \varepsilon)$ $X^{(k+1)}; y^{(k+1)}$ et $Z^{(k+1)}$ ne sont pas les solutions optimales approximatives de (2.2), dans ce cas, le paramètre μ se réduit à $\mu^+ = (1 - \theta)\mu$ avec $0 < \theta < 1$, et restons proche de la trajectoire centrale. Généralement, si θ dépend de n, notamment $\theta = O(\frac{1}{\sqrt{n}})$, alors l'algorithme est appelé : à petit pas "short-step", et si $\theta = O(1)$ est un constant, alors l'algorithme est appelé : à grand pas "long-step".

2.4.2 Les directions de Newton

Soient $\mu > 0$ et $(X; y; Z)$ un point strictement réalisable primal-dual (mais $XZ \neq \mu I$), la direction de Newton $(\Delta X; \Delta y; \Delta Z)$ en ce point est la solution unique du système linéaire suivant :

$$\begin{cases} A_i \bullet \Delta X &= 0; \; i = 1; \dots; m; \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) &= 0; \\ \Delta X Z + X \Delta Z &= \mu I - XZ; \end{cases} \tag{2.3}$$

qui est équivalent à :

$$\begin{cases} A_i \bullet \Delta X = 0, \quad i = 1, \ldots, m; \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0; \\ \Delta X + X \Delta Z Z^{-1} = \mu Z^{-1} - X. \end{cases} \tag{2.4}$$

Sous les hypothèses précédentes ce système linéaire admet une solution unique $(\Delta X; \Delta y; \Delta Z)$.

On note que ΔZ est symétrique d'après la deuxième équation du système (2.4) mais ΔX n'est pas forcément symétrique. Alors pour traiter ce problème plusieurs travaux ont été proposés (voir [7], [8], [23]), où l'idée de toutes les propositions est d'introduire une matrice inversible P (scaling matrix) et de remplacer le système (2.4) par le système suivant :

$$\begin{cases} A_i \bullet \Delta X = 0, \quad i = 1, \ldots, m; \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0; \\ \Delta X + P \Delta Z P^{T} = \mu Z^{-1} - X. \end{cases} \tag{2.5}$$

L'un des choix de la matrice P est :

$$P := X^{\frac{1}{2}}(X^{\frac{1}{2}} Z X^{\frac{1}{2}})^{-\frac{1}{2}} X^{\frac{1}{2}}, \text{ (Nesterov Todd scaling matrix)},$$

où $P \in S^n_{++}$.

Maintenant, on définit la matrice symétrique et définie positive $D = P^{\frac{1}{2}}$. Le rôle de D est de mettre à l'échelle les deux matrices X et Z au même matrice symétrique et définie positive V qui vérifie :

$$V = \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} D Z D = \frac{1}{\sqrt{\mu}} (D^{-1} X Z D)^{1/2}. \tag{2.6}$$

De plus, les nouvelles directions (scaling directions) D_X et D_Z sont définies par :

$$D_X := \frac{1}{\sqrt{\mu}} D^{-1} \Delta X D^{-1}; \quad D_Z := \frac{1}{\sqrt{\mu}} D \Delta Z D. \tag{2.7}$$

Posons $(D_y)_i = \Delta y_i$; alors le système (2.5) devient :

$$\begin{cases} \tilde{A}_i \bullet D_X & = 0;\ i = 1; \dots ; m; \\ \sum_{i=1}^{m} (D_y)_i \tilde{A}_i + D_Z - \tilde{Q}(D_X) & = 0; \\ D_X + D_Z & = P_V; \end{cases} \tag{2.8}$$

où

$$P_V = V^{-1} - V; \ \tilde{A}_i = \frac{1}{\sqrt{\mu}} D A_i D;\ i = 1; \dots ; m; \text{et } \tilde{Q}(D_X) = D Q(D D_X D) D. \tag{2.9}$$

On note que ce dernier système est équivalent au système (2.5), les matrices $[\tilde{A}_1; \tilde{A}_2; \dots; \tilde{A}_m]$ sont aussi linéairement indépendantes et l'opérateur $\tilde{Q}$ est aussi auto-adjoint et monotone. De plus, les directions D_X, D_Z sont symètriques et vérifiants ([17], [18], [29]) :

$$D_X \bullet D_Z = \frac{1}{\mu} \Delta X \bullet Q(\Delta X) \geq 0; \tag{2.10}$$

car

$$\begin{aligned} D_X \bullet D_Z &= \frac{1}{\sqrt{\mu}}(D^{-1} \Delta X D^{-1}) \bullet \frac{1}{\sqrt{\mu}}(D \Delta Z D);\ D = D^{\top}; \\ &= \frac{1}{\mu}(D D^{-1} \Delta X D^{-1} D) \bullet \Delta Z; \\ &= \frac{1}{\mu} \Delta X \bullet \Delta Z; \end{aligned}$$

et on a :

$$\sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0 \Longrightarrow \Delta Z = -\sum_{i=1}^{m} \Delta y_i A_i + Q(\Delta X);$$

alors

$$\begin{aligned} D_X \bullet D_Z &= \frac{1}{\mu} < \Delta X; -\sum_{i=1}^{m} \Delta y_i A_i + Q(\Delta X) >; \\ &= \frac{1}{\mu} \left[-\sum_{i=1}^{m} \Delta y_i (A_i \bullet \Delta X) + \Delta X \bullet Q(\Delta X) \right]; \text{ où } A_i \bullet \Delta X = 0; \\ &= \frac{1}{\mu} \Delta X \bullet Q(\Delta X) \geq 0; \text{ car } Q \text{ est monotone.} \end{aligned}$$

Notons que les directions cherchées ne sont pas orthogonales contrairement au cas (SDP), ce qui fait l'analyse théorique de la complexité de l'algorithme sera difficile.

2.4.3 La mesure de proximité

Pour que les itérations produites par l'algorithme restent réalisables et proches de la trajectoire centrale (TC), on est obligé d'introduire une mesure de proximité. Pour cela on définit :

$$\delta := \delta(X; Z; \mu) = \frac{\|P_V\|_F}{2} = \frac{1}{2}\left\|V^{-1} - V\right\|_F .$$

C'est facile de vérifier que :

$$\delta(X; Z; \mu) = 0 \iff V^{-1} = V \iff XZ = \mu I;$$

qui justifie l'appartenance des points à la trajectoire centrale.
Dans cet algorithme, on utilise une valeur seuil β de la proximité et on suppose que $(X^0; y^0; Z^0)$ est un point strictement réalisable avec $\delta(X^0; Z^0; \mu^0) \leq \beta$ pour certain μ^0 connu.

Maintenant, on présente un algorithme primal-dual à petit pas pour tracer approximativement la trajectoire centrale.

2.4.4 Algorithme

Dans cette partie, on présente un algorithme de trajectoire centrale de type primal-dual pour (QSDP) :

Début d'algorithme
Données :
un paramètre de précision $\varepsilon > 0$;
un paramètre de proximité $0 < \beta < 1$ (défaut $\beta = \frac{1}{\sqrt{2}}$);
un paramètre $0 < \theta < 1$ (défaut $\theta = \frac{1}{2\sqrt{n}}$);
un paramètre de barrière μ^0 (fixé).
Initilisation : soit $(X^0; y^0; Z^0)$ vérifie la CPI tel que $\delta(X^0; Z^0; \mu^0) \leq \beta$ et $k = 0$;
Tant que $n\mu^{(k)} \geq \varepsilon$ faire :
• $\mu^{(k+1)} = (1 - \theta)\mu^{(k)}$.
• calculer $(\Delta X^{(k)}; \Delta y^{(k)}; \Delta Z^{(k)})$ via le système (2.8);
• mise à jour $X^{(k+1)} = X^{(k)} + \Delta X^{(k)}$; $y^{(k+1)} = y^{(k)} + \Delta y^{(k)}$; $Z^{(k+1)} = Z^{(k)} + \Delta Z^{(k)}$ et $k = k + 1$;
Fin tant que.
Fin algorithme.

Algorithme 2.4.4

Le choix des paramètres β et θ décrit dans l'Algorithme 2.4.4 sera justifié par la suite.

2.4.5 La convergence de l'algorithme et l'analyse de la complexité

On considère la matice symétrique suivante :

$$D_{XZ} = \frac{1}{2}[D_X D_Z + D_Z D_X];$$

qui sera utile par la suite et $0 \leq \alpha \leq 1$.

Maintenant, pour étudier la convergence de l'algorithme et l'analyse de la complexité, on a besoin de quelques résultats techniques.

Lemme 2.4.4 *[7, Lemme 3.3.1] Soient* $X(\alpha) := X + \alpha\Delta X$; $Z(\alpha) := Z + \alpha\Delta Z$ *telles que* $X \succ 0$ *et* $Z \succ 0$. *Si*

$$\det(X(\alpha)Z(\alpha)) > 0; \forall 0 \leq \alpha \leq \bar{\alpha}.$$

Alors : $X(\alpha) \succ 0$ *et* $Z(\alpha) \succ 0$.

Lemme 2.4.5 *[7, Lemme 3.3.3] Soient* $Q \in S^n$ *et* $M \in \mathbb{R}^{n \times n}$ *une matrice anti-symétrique. Si* $Q \succ 0$ *alors* $\det(Q + M) > 0$. *De plus, si on* $\lambda_i(Q + M) \in \mathbb{R}$ *et* $i = 1, \ldots, n$, *alors*

$$0 < \lambda_{min}(Q) \leq \lambda_{min}(Q + M) \leq \lambda_{max}(Q + M) \leq \lambda_{max}(Q).$$

Lemme 2.4.6 *Soit* (D_X, D_Z) *une solution du système (2.8) et* $\mu > 0$. *Si* $\delta := \delta(X, Z, \mu)$ *alors*

$$0 \leq D_X \bullet D_Z \leq 2\delta^2. \tag{2.11}$$

De plus, la norme spectrale de D_{XZ} *vérifie :*

$$\|D_{XZ}\|_2 \leq \frac{1}{4}\|D_X + D_Z\|_F^2 = \delta^2. \tag{2.12}$$

Pour montrer ce lemme, on présente la remarque suivante :

Remarque 2.4.1 *En (QSDP), on sait que les directions* D_X *et* D_Z *ne sont pas orthogonales c-à-d* $D_X \bullet D_Z \geq 0$ *ce qui implique :*

$$\|D_X + D_Z\|_F \geq \|D_X - D_Z\|_F.$$

Preuve : (du lemme [2.4.6]) D'une part on a :

$$D_X \bullet D_Z = \frac{1}{\mu} \Delta X \bullet Q(\Delta X) \geq 0. \tag{2.13}$$

D'autre part on a :

$$\begin{aligned} 2\delta^2 &= \frac{1}{2}\|P_V\|_F^2 = \frac{1}{2}\|D_X + D_Z\|_F^2, \\ &= \frac{1}{2}\mathrm{Tr}\left((D_X + D_Z)^2\right), \\ &= \frac{1}{2}\left[\|D_X\|_F^2 + \|D_Z\|_F^2 + 2D_X \bullet D_Z\right], \\ &= \frac{1}{2}\left[\|D_X\|_F^2 + \|D_Z\|_F^2\right] + D_X \bullet D_Z \geq D_X \bullet D_Z, \end{aligned} \tag{2.14}$$

de (2.13) et (2.14) on obtient :

$$0 \leq D_X \bullet D_Z \leq 2\delta^2.$$

Maintenant, on montre que :

$$\|D_{XZ}\|_2 \le \frac{1}{4}\|D_X + D_Z\|_F^2 = \delta^2.$$

Par définition on a :

$$\delta = \frac{1}{2}\|P_V\|_F = \frac{1}{2}\|D_X + D_Z\|_F ;$$

alors

$$\frac{1}{4}\|D_X + D_Z\|_F^2 = \delta^2; \tag{2.15}$$

il nous reste à montrer que :

$$\|D_{XZ}\|_2 \le \frac{1}{4}\|D_X + D_Z\|_F^2 .$$

On a :

$$D_X D_Z + D_Z D_X = \frac{1}{2}\left[(D_X + D_Z)^2 - (D_X - D_Z)^2\right];$$

ce qui implique que :

$$-\frac{1}{4}(D_X - D_Z)^2 \preceq D_{XZ} \preceq \frac{1}{4}(D_X + D_Z)^2;$$

et que :

$$-\frac{1}{4}\|D_X - D_Z\|_F^2 I \preceq D_{XZ} \preceq \frac{1}{4}\|D_X + D_Z\|_F^2 I.$$

D'après la Remarque [2.4.1], on a :

$$\|D_X + D_Z\|_F \ge \|D_X - D_Z\|_F ;$$

qui est équivalent à :

$$-\frac{1}{4}\|D_X + D_Z\|_F^2 \le -\frac{1}{4}\|D_X - D_Z\|_F^2 ,$$

donc

$$-\frac{1}{4}\|D_X + D_Z\|_F^2 I \preceq D_{XZ} \preceq \frac{1}{4}\|D_X + D_Z\|_F^2 I;$$

en utilisant les valeurs propres, on obtient :

$$-\frac{1}{4}\|D_X + D_Z\|_F^2 \le \lambda_i(D_{XZ}) \le \frac{1}{4}\|D_X + D_Z\|_F^2\,;\forall i = 1,\dots,n.$$

Alors

$$|\lambda_i(D_{XZ})| \le \frac{1}{4}\|D_X + D_Z\|_F^2\,;\forall i = 1,\dots,n.$$

Et par conséquent

$$\max_{i=1}^{n} |\lambda_i(D_{XZ})| = \rho(D_{XZ}) \le \frac{1}{4}\|D_X + D_Z\|_F^2\,;$$

puisque D_{XZ} est symétrique, alors :

$$\rho(D_{XZ}) = \|D_{XZ}\|_2 \le \frac{1}{4}\|D_X + D_Z\|_F^2 \tag{2.16}$$

De (2.15) et (2.16) on déduit que :

$$\|D_{XZ}\|_2 \le \frac{1}{4}\|D_X + D_Z\|_F^2 = \delta^2.$$

Ce qui achève la preuve. ■

Lemme 2.4.7 *On a :*

$$\|D_{XZ}\|_F^2 \le \frac{1}{8}\|P_V\|_F^4\,.$$

Preuve : On sait que : $D_X D_Z + D_Z D_X = \frac{1}{2}[(D_X + D_Z)^2 - (D_X - D_Z)^2]$. Afin de simplifier les calculs, on pose

$$P_V = D_X + D_Z;\ \text{et}\ Q_V = D_X - D_Z;$$

alors

$$D_{XZ} = \frac{1}{4}\left[P_V^2 - Q_V^2\right];$$

en utilisant la norme de Frobenius, on obtient :

$$\begin{aligned} \|D_{XZ}\|_F^2 &= \left\| \frac{1}{4}(P_V^2 - Q_V^2) \right\|_F^2 = \frac{1}{16} \left\| P_V^2 - Q_V^2 \right\|_F^2, \\ &\leq \frac{1}{16} \left[\left\| P_V^2 \right\|_F^2 + \left\| Q_V^2 \right\|_F^2 \right], \end{aligned}$$

puisque

$$\left\| P_V^2 \right\|_F^2 \leq \|P_V\|_F^2 \, \|P_V\|_F^2 = \|P_V\|_F^4,$$

alors

$$\|D_{XZ}\|_F^2 \leq \frac{1}{16} \left[\|P_V\|_F^4 + \|Q_V\|_F^4 \right],$$

D'après la Remarque [2.4.1], on a :

$$\|D_X + D_Z\|_F \geq \|D_X - D_Z\|_F \Leftrightarrow \|P_V\|_F \geq \|Q_V\|_F.$$

Donc

$$\|D_{XZ}\|_F^2 \leq \frac{1}{16} \left[\|P_V\|_F^4 + \|P_V\|_F^4 \right] = \frac{1}{8} \|P_V\|_F^4.$$

Ce qui achève la preuve. ■

Dans les lemmes suivants, on va analyser la complexité de l'algorithme, on commence de donner quelques conditions pour assurer la stricte faisabilité du pas de Newton complète.

Analyse de la stricte faisabilité

Lemme 2.4.8 *(Condition suffisante) Si*

$$\delta := \delta(X, Z, \mu) < 1,$$

alors le pas de Newton complète est strictement réalisable c-à-d :

$$X(1) \succ 0 \text{ et } Z(1) \succ 0.$$

Preuve : D'après le Lemme [2.4.4], il suffit de montrer que $\det(X(\alpha)Z(\alpha))$ est positif pour tout $0 \leq \alpha \leq 1$; alors $X(1) \succ 0$ et $Z(1) \succ 0$. Pour obtenir ce résultat, on note que :

$$\begin{aligned} X(\alpha)Z(\alpha) &= (X + \alpha\Delta X)(Z + \alpha\Delta Z); \\ &= XZ + \alpha(\Delta XZ + X\Delta Z) + \alpha^2 \Delta X \Delta Z. \end{aligned}$$

D'après (2.6) et (2.7), on obtient :

$$\begin{aligned} X(\alpha)Z(\alpha) &= \mu DV^2D^{-1} + \alpha(\mu DD_XVD^{-1} + \mu DVD_ZD^{-1}) + \alpha^2\mu DD_XD_ZD^{-1}; \\ &= \mu D\left[V^2 + \alpha(D_XV + VD_Z) + \alpha^2 D_XD_Z\right]D^{-1}; \\ &\sim \mu\left[V^2 + \alpha(D_XV + VD_Z) + \alpha^2 D_XD_Z\right]; \\ &= \mu\left[V^2 + \alpha(D_XV + VD_Z) + \alpha^2 D_XD_Z\right] + \frac{1}{2}\mu\alpha^2 D_ZD_X - \frac{1}{2}\mu\alpha^2 D_ZD_X \\ &\quad + \frac{1}{2}\mu\alpha(VD_X + D_ZV) - \frac{1}{2}\mu\alpha(VD_X + D_ZV); \\ &= B(\alpha) + M(\alpha); \end{aligned} \tag{2.17}$$

avec

$$B(\alpha) = \mu\left(V^2 + \frac{1}{2}\alpha(D_XV + VD_Z + VD_X + D_ZV) + \frac{1}{2}\alpha^2(D_XD_Z + D_ZD_X)\right);$$

et

$$M(\alpha) = \mu\left(\frac{1}{2}\alpha(D_XV + VD_Z - VD_X - D_ZV) + \frac{1}{2}\alpha^2(D_XD_Z - D_ZD_X)\right).$$

Où $B(\alpha)$ est une matrice symétrique et $M(\alpha)$ est anti-symétrique. Alors d'après le Lemme [2.4.5], on a :

$$\det(X(\alpha)Z(\alpha)) > 0;$$

si la matrice $B(\alpha)$ est définie positive, $\forall 0 \leq \alpha \leq 1$. Pour cela, on écrit la matrice $B(\alpha)$ sous la forme :

$$\begin{aligned} B(\alpha) &= \mu\left(V^2 + \frac{1}{2}\alpha((D_X + D_Z)V + V(D_Z + D_X)) + \alpha^2 D_{XZ}\right); \text{où } D_X + D_Z = P_V; \\ &= \mu\left(V^2 + \frac{1}{2}\alpha(P_VV + VP_V) + \alpha^2 D_{XZ}\right); \end{aligned}$$

donc d'après (2.9), $B(\alpha)$ devient :

$$\begin{aligned} B(\alpha) &= \mu\left[V^2 + \frac{1}{2}\alpha(V^{-1}V - V^2 + VV^{-1} - V^2) + \alpha^2 D_{XZ}\right]; \\ &= \mu\left[(1-\alpha)V^2 + \alpha(I + \alpha D_{XZ})\right]; \end{aligned} \tag{2.18}$$

et par hypothèse on a $\delta < 1$; alors d'après le Lemme [2.4.6], on déduit que

$$\|D_{XZ}\|_2 < 1;$$

et par conséquent $\|\alpha D_{XZ}\|_2 < 1$; pour tout $0 \leq \alpha \leq 1$. Cela implique que

$$(I + \alpha D_{XZ}) \in S^n_{++}; \forall 0 \leq \alpha \leq 1.$$

D'autre part on sait que $V \in S^n_{++}$ et $0 \leq \alpha \leq 1$; donc

$$(1-\alpha)V^2 \in S^n_{++}.$$

Alors, $B(\alpha)$ est définie positive pour tout $0 \leq \alpha \leq 1$ si $\delta < 1$. De plus, puisque $X(0) \succ 0$ et $Z(0) \succ 0$; le Lemme [2.4.4] implique que $X(1) \succ 0$ et $Z(1) \succ 0$; ce qui achève la preuve. ■

Pour simplifier, on peut écrire $X_+ = X(1)$ et $Z_+ = Z(1)$. Alors :

$$X_+Z_+ \sim \mu V_+^2.$$

La convergence quadratique de la mesure de proximité

Pour montrer la convergence quadratique de la mesure de proximité, on a besoin d'un résultat technique concernant le $sp(V_+^2)$.

Lemme 2.4.9 *On a :*

$$\lambda_{\min}(V_+^2) \geq (1-\delta^2).$$

Où $\lambda_{\min}$ *est la plus petite valeur propre de* V_+^2.

Preuve : D'après (2.17) dans le Lemme [2.4.8], on pose $\alpha = 1$; on obtient :

$$\begin{aligned} X(1)Z(1) &\sim B(1) + M(1); \\ &\sim \mu I + \mu D_{XZ} + M(1) = \mu(I + D_{XZ}) + M(1); \end{aligned}$$

Comme $X_+Z_+ \sim \mu V_+^2$, alors :

$$\lambda_{\min}(V_+^2) = \lambda_{\min}\left((I + D_{XZ}) + \frac{1}{\mu} M(1)\right);$$

où $(I + D_{XZ}) \in S^n$ et $M(1) = -M(1)^T$. Alors d'après le Lemme [2.4.5], on obtient :

$$\begin{aligned}
\lambda_{\min}(V_+^2) &\geq \lambda_{\min}(I + D_{XZ}), \\
&\geq 1 - |\lambda_{\min}(D_{XZ})|; \\
&\geq 1 - \rho(D_{XZ}) = 1 - \|D_{XZ}\|_2; \\
&\geq 1 - \frac{1}{4}\|P_V\|_F^2; \text{ d'après le Lemme [2.4.6].} \\
&= (1 - \delta^2).
\end{aligned}$$

Ce qui achève la preuve. ■

Maintenant, on montre la convergence quadratique de la mesure de proximité.

Lemme 2.4.10 *Si $\delta := \delta(X; Z; \mu) < 1$, alors :*

$$\delta_+ := \delta(X_+; Z_+; \mu) \leq \frac{\delta^2}{\sqrt{2(1-\delta^2)}}.$$

Preuve : On a :

$$\begin{aligned}
\delta_+^2 &= \frac{1}{4}\left\|V_+^{-1} - V_+\right\|_F^2 = \frac{1}{4}\left\|V_+^{-1}(I - V_+^2)\right\|_F^2; \\
&\leq \frac{1}{4}\lambda_{\max}^2(V_+^{-1})\left\|I - V_+^2\right\|_F^2; \text{ d'après le Lemme [1.2.2],} \\
&= \frac{1}{4}\frac{1}{\lambda_{\min}^2(V_+)}\left\|I - V_+^2\right\|_F^2; \\
&= \frac{1}{4}\frac{1}{\lambda_{\min}(V_+^2)}\left\|I - V_+^2\right\|_F^2.
\end{aligned}$$

D'après le Lemme [2.4.9], on a :

$$\delta_+^2 \leq \frac{1}{4(1-\delta^2)}\left\|I - V_+^2\right\|_F^2.$$

Pour compléter la démonstration, on montre que :

$$\left\|I - V_+^2\right\|_F^2 \leq \|D_{XZ}\|_F^2.$$

Puisque $X(1)Z(1) \preceq \mu V_+^2$ et d'après le Lemme [2.4.8], on déduit que :

$$
\begin{aligned}
I - V_+^2 &= I - \frac{1}{\mu} D^{-1} X(1)Z(1) D; \\
&= I - \frac{1}{\mu} D^{-1} X_+ Z_+ D; \\
&= I - \frac{1}{\mu} D^{-1} [B(1) + M(1)] D; \\
&\preceq I - \frac{1}{\mu} B(1) - \frac{1}{\mu} M(1);
\end{aligned}
$$

d'après l'expression (2.18) de Lemme [2.4.8] avec $\alpha = 1$; on obtient :

$$B(1) = \mu (I + D_{XZ});$$

alors

$$I - V_+^2 \preceq -D_{XZ} - \frac{1}{\mu} M(1).$$

En utilisant la norme de Frobenius, on obtient :

$$
\begin{aligned}
\left\| I - V_+^2 \right\|_F^2 &= \sum_i \left[\lambda_i\left(-D_{XZ} - \frac{1}{\mu} M(1)\right)\right]^2; \\
&= \operatorname{Tr}\left(D_{XZ} + \frac{1}{\mu} M(1)\right)^2; \\
&= \operatorname{Tr}\left(D_{XZ}^2 - \frac{1}{\mu^2} M(1)M(1)^T + \frac{2}{\mu}(D_{XZ} M(1)\right);
\end{aligned}
$$

puisque $M(1) = -M(1)^T$ et $D_{XZ} \in S^n$, alors $(D_{XZ} M(1))$ est anti-symétrique ce qui implique que

$$
\begin{aligned}
\left\| I - V_+^2 \right\|_F^2 &= \operatorname{Tr}\left(D_{XZ}^2 - \frac{1}{\mu^2} M(1)M(1)^T\right); \text{ car } \operatorname{Tr}(D_{XZ} M(1)) = 0; \\
&\leq \operatorname{Tr}(D_{XZ}^2) = \|D_{XZ}\|_F^2 .
\end{aligned}
$$

Et par conséquent,

$$\left\| I - V_+^2 \right\|_F^2 \leq \|D_{XZ}\|_F^2 ;$$

d'après le Lemme [2.4.7], on déduit que :

$$\delta_+^2 \leq \frac{1}{4(1-\delta^2)} \|D_{XZ}\|_F^2 \leq \frac{1}{4(1-\delta^2)} \frac{1}{8} \|P_V\|_F^4 \, ;$$

c'est équivalent à :

$$\delta_+^2 \leq \frac{1}{2(1-\delta^2)} \delta^4; \text{ où } \delta^4 = \frac{1}{16} \|P_V\|_F^4 \, .$$

Finalement,

$$\delta_+ := \delta(X_+; Z_+; \mu) \leq \frac{\delta^2}{\sqrt{2(1-\delta^2)}}.$$

Ce qui achève la preuve. ■

Le corollaire suivant montre la convergence quadratique de la mesure de proximité avec le pas de Newton complète.

Corollaire 2.4.1 *Si* $\delta(X; Z; \mu) < \frac{1}{\sqrt{2}}$*, alors* $\delta_+ < \delta^2(X; Z; \mu)$*; c-à-d : on obtient la convergence quadratique.*

L'influence de nouveau itéré en saut de dualité

Dans le lemme suivant, on analyse l'influence du pas complète de Newton sur le saut de dualité.

Lemme 2.4.11 *Après le pas de Newton complète et si* $\delta(X; Z; \mu) < \frac{1}{\sqrt{2}}$*; alors le saut de dualité vérifie :*

$$X_+ \bullet Z_+ \leq \mu(n+1). \tag{2.19}$$

Preuve : on a :

$$\begin{aligned}
X_+ \bullet Z_+ &= \mu \, Tr(V_+^2) = Tr(B(1) + M(1)) = Tr(\mu(I + D_{XZ}) + M(1)); \\
&= \mu \, Tr((I + D_{XZ}) + \frac{1}{\mu} M(1)); \\
&= \mu \, Tr(I + D_{XZ}); \text{ car si } M(1) = -M(1)^T \Rightarrow Tr(M(1)) = 0. \\
&= \mu[n + Tr(D_{XZ})]; \\
&= \mu[n + \frac{1}{2} Tr(D_X D_Z + D_Z D_X)]; \\
&= \mu[n + Tr(D_X D_Z)]; \text{car } D_X \bullet D_Z = D_Z \bullet D_X; \\
&= \mu[n + D_X \bullet D_Z].
\end{aligned}$$

D'après l'expression (2.11) du Lemme [2.4.6], et $\delta < \frac{1}{\sqrt{2}}$; on obtient :

$$D_X \bullet D_Z \leq 2\left(\frac{1}{\sqrt{2}}\right)^2 = 1;$$

par conséquent,

$$X_+ \bullet Z_+ \leq \mu(n + 1).$$

Ce qui achève la preuve. ■

La mise à jour du paramètre barrière μ

La recherche d'une solution optimale primale-duale est équivalente à mettre le saut de dualité $X \bullet Z$ vers le zéro, cela est exprimé par la mise à jour du paramètre μ via :

$$\mu_+ = (1 - \theta)\mu; \; 0 < \theta < 1.$$

Par la suite on montre l'influence de la mise à jour du μ sur la nouvelle proximité après un pas complète de Newton.

Lemme 2.4.12 *Si* $\delta(X; Z; \mu) < \frac{1}{\sqrt{2}}$ *et* $\mu_+ = (1 - \theta)\mu$*, où* $0 < \theta < 1$. *Alors*

$$\delta^2(X_+; Z_+; \mu_+) \leq (1 - \theta)\delta_+^2 + \frac{\theta^2(n + 1)}{4(1 - \theta)} + \frac{\theta}{2}. \tag{2.20}$$

De plus, si $\delta := \delta(X;Z;\mu) < \frac{1}{\sqrt{2}}, \theta = \frac{1}{2\sqrt{n}}$ *et* $n \geq 2$, *alors* :

$$\delta^2(X_+;Z_+;\mu_+) < \frac{1}{2}.$$

Preuve : Soient $V_+^2 = \frac{1}{\mu} D^{-1} X_+ Z_+ D$; $\mu_+ = (1-\theta)\mu$ et $\delta < \frac{1}{\sqrt{2}}$. Alors,

$$\begin{aligned}
4\delta^2(X_+;Z_+;\mu_+) &= \left\| \sqrt{1-\theta}\, V_+^{-1} - \frac{1}{\sqrt{1-\theta}} V_+ \right\|_F^2; \\
&= \left\| \sqrt{1-\theta}\, V_+^{-1} - \frac{(1-\theta)+\theta}{\sqrt{1-\theta}} V_+ \right\|_F^2; \\
&= \left\| \sqrt{1-\theta}\,(V_+^{-1} - V_+) - \frac{\theta}{\sqrt{1-\theta}} V_+ \right\|_F^2; \\
&= (1-\theta) \left\| V_+^{-1} - V_+ \right\|_F^2 + \frac{\theta^2}{1-\theta} \|V_+\|_F^2 - 2\theta \operatorname{Tr}\left((V_+^{-1} - V_+)V_+\right); \\
&= (1-\theta) \left\| V_+^{-1} - V_+ \right\|_F^2 + \frac{\theta^2}{1-\theta} \|V_+\|_F^2 - 2\theta \operatorname{Tr}\left(I - V_+^2\right); \\
&= (1-\theta)4\delta_+^2 + \frac{\theta^2}{1-\theta} \|V_+\|_F^2 - 2\theta n + 2\theta \|V_+\|_F^2;
\end{aligned}$$

d'après le Lemme [2.4.11], on a :

$$\|V_+\|_F^2 = \operatorname{Tr}(V_+^2) = \frac{1}{\mu} X_+ \bullet Z_+ \leq (n+1).$$

Donc

$$4\delta^2(X_+;Z_+;\mu_+) \leq (1-\theta)4\delta_+^2 + \frac{\theta^2(n+1)}{1-\theta} - 2\theta n + 2\theta(n+1);$$

par conséquent

$$\delta^2(X_+;Z_+;\mu_+) \leq (1-\theta)\delta_+^2 + \frac{\theta^2(n+1)}{4(1-\theta)} + \frac{\theta}{2}.$$

Comme $\delta < \frac{1}{\sqrt{2}}$; d'après le Corollaire [2.4.1] on obtient $\delta_+^2 < \delta^4 < \frac{1}{4}$; et par conséquent

$$\delta^2(X_+;Z_+;\mu_+) \leq \frac{(1-\theta)}{4} + \frac{\theta^2(n+1)}{4(1-\theta)} + \frac{\theta}{2}.$$

Maintenant, on considère $\theta = \frac{1}{2\sqrt{n}}$; alors

$$\begin{aligned}\delta^2(X_+, Z_+, \mu_+) &\leq \frac{(1-\theta)}{4} + \frac{\frac{1}{4n}(n+1)}{4(1-\theta)} + \frac{\theta}{2}, \quad \frac{n+1}{4n} \leq \frac{3}{8}, \forall n \geq 2,\\ &\leq \frac{(1-\theta)}{4} + \frac{\theta}{2} + \frac{3}{32(1-\theta)},\end{aligned}$$

pour $n \geq 2$, on a: $0 \leq \theta \leq \frac{1}{\sqrt{8}}$ et la fonction

$$f(\theta) = \frac{(1-\theta)}{4} + \frac{\theta}{2} + \frac{3}{32(1-\theta)},$$

est strictement croissante sur l'intervalle $[0, \frac{1}{\sqrt{8}}]$. Par conséquent,

$$f(\theta) \leq f\left(\frac{1}{\sqrt{8}}\right) \approx 0.48341 < \frac{1}{2}, \forall \theta \in \left[0, \frac{1}{\sqrt{8}}\right],$$

finalement $\delta^2(X_+, Z_+, \mu_+) < \frac{1}{2}$. Ce qui achève la preuve. ■

Analyse de la complexité

Dans cette partie, on présente le nombre des itérations à partir duquel l'Algorithme 2.4.4 converge.

Lemme 2.4.13 *Soient X^{k+1} et Z^{k+1} le (k+1)ième itération produite par l'Algorithme 2.4.4, avec $\mu := \mu_k$, alors*

$$X^{k+1} \bullet Z^{k+1} \leq \varepsilon,$$

est satisfaite si

$$k \geq \left\lceil \frac{1}{\theta} \log\left(\frac{\mu_0 n}{\varepsilon}\right) \right\rceil.$$

Preuve : D'après le Lemme [2.4.11], on a :

$$X^{k+1} \bullet Z^{k+1} \leq \mu_k(n+1), \tag{2.21}$$

avec

$$\mu_k = (1-\theta)\mu_{k-1} = (1-\theta)^k \mu_0.$$

Alors (2.21) devient :

$$X^{k+1} \bullet Z^{k+1} \leq (1-\theta)^k \mu_0 (n+1);$$

où l'inégalité $X^{k+1} \bullet Z^{k+1} \leq \varepsilon$ est satisfaite si :

$$(1-\theta)^k \mu_0 (n+1) \leq \varepsilon.$$

En utilisant la fonction croissante log on obtient :

$$k \log(1-\theta) \leq \log \varepsilon - \log(\mu_0 (n+1)) \leq \log \varepsilon - \log(\mu_0 n);$$

puisque $\log(1-\theta) \leq -\theta$, pour $0 < \theta < 1$; alors

$$k\theta \geq \log\left(\frac{\mu_0 n}{\varepsilon}\right).$$

Par conséquent

$$k \geq \frac{1}{\theta} \log\left(\frac{\mu_0 n}{\varepsilon}\right).$$

Ce qui achève la preuve. ∎

On note que pour $\theta = \frac{1}{2\sqrt{n}}$; on obtient le théorème suivant :

Théorème 2.4.1 *Soient* $\theta = \frac{1}{2\sqrt{n}}$, $\beta = \frac{1}{\sqrt{2}}$. *Alors, l'algorithme de la trajectoire centrale de type primal-dual avec le pas de Newton complète nécessite au plus de* $O(\sqrt{n} \log(\frac{n\mu_0}{\varepsilon}))$ *itérations.*

Chapitre 3

Implémentation numérique

Dans ce chapitre, on s'intéresse au calcul de la direction de déplacement où on donne une condition suffisante pour assurer l'existence et l'unicité de la direction de NT, calcul de la matrice P et ainsi V.

3.1 Calcul de la direction

On a vu dans le Chapitre 2 que la direction de NT est donnée par le système linéaire suivant :

$$\begin{cases} A_i \bullet \Delta X = 0, \; i = 1, \ldots, m; \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0; \\ \Delta X + P \Delta Z P^{T} = \mu Z^{-1} - X; \end{cases}$$

qui s'écrit sous la forme :

$$\begin{cases} A_i \bullet \Delta X = 0, \; i = 1, \ldots, m; \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0; \\ (Z \Delta X + \Delta X Z) + (Z P \Delta Z P^{T} + P \Delta Z P^{T} Z) = 2\mu I_n - (XZ + ZX). \end{cases}$$

En appliquant les propriétés de l'opérateur "svec" sur la première équation, on obtient :

$$
\begin{aligned}
A_i \bullet \Delta X = 0; \forall i = 1;\dots;m; \quad &\Leftrightarrow \quad \operatorname{svec}^T(A_i)\operatorname{svec}(\Delta X) = 0; \forall i = 1;\dots;m;\\
&\Leftrightarrow \quad \begin{bmatrix} \operatorname{svec}^T(A_1) \\ \vdots \\ \operatorname{svec}^T(A_m) \end{bmatrix} \operatorname{svec}(\Delta X) = 0;\\
&\Leftrightarrow \quad A \operatorname{svec}(\Delta X) = 0;
\end{aligned}
$$

où $A \in \Re^{m \times \bar{n}}$ et $\operatorname{svec}(\Delta X) \in \Re^{\bar{n}}$.
La deuxième équation du système précédent devient :

$$A^T \Delta y + \operatorname{svec}(\Delta Z) - \tilde{Q} \operatorname{svec}(\Delta X) = 0;$$

où $\tilde{Q} \in S^{\bar{n}}$ sera définie dans le chapitre suivant,

$$A^T = \begin{bmatrix} \operatorname{svec}(A_1) & \operatorname{svec}(A_2) & \dots & \operatorname{svec}(A_m) \end{bmatrix} \text{ et } A^T \in \Re^{\bar{n} \times m};$$

car :

$$
\begin{aligned}
A^T \Delta y &= \begin{bmatrix} \operatorname{svec}(A_1) & \dots & \operatorname{svec}(A_m) \end{bmatrix} \Delta y;\\
&= \sum_{i=1}^{m} \operatorname{svec}(A_i)(\Delta y)_i;
\end{aligned}
$$

et

$$
\begin{aligned}
\operatorname{svec}(Q(\Delta X)) &= \operatorname{svec}(Q(\Delta X)I_n^T);\\
&= \frac{1}{2}\operatorname{svec}(Q(\Delta X)I_n^T + I_n Q(\Delta X));\\
&= \tilde{Q} \operatorname{svec}(\Delta X); \text{ voir [26].}
\end{aligned}
$$

La troisième équation devient :

$$E \operatorname{svec}(\Delta X) + F \operatorname{svec}(\Delta Z) = w;$$

où

$$E = (Z \otimes_s I_n); \; F = (ZP \otimes_s P) \; \text{ et } \; w = \operatorname{svec}(\mu I_n - \frac{1}{2}(XZ + ZX)).$$

Finalement, on obtient le système d'équations linéaires suivant :

$$\begin{cases} A\,\mathrm{svec}(\Delta X) & = 0; \\ A^T \Delta y - Q\,\mathrm{svec}(\Delta X) + I_{\tilde{n}}\,\mathrm{svec}(\Delta Z) & = 0; \\ E\,\mathrm{svec}(\Delta X) + F\,\mathrm{svec}(\Delta Z) & = w; \end{cases}$$

qui s'écrit sous la forme matricielle suivante :

$$\begin{bmatrix} 0 & A & 0 \\ A^T & -Q & I_{\tilde{n}} \\ 0 & E & F \end{bmatrix} \begin{bmatrix} \Delta y \\ \mathrm{svec}(\Delta X) \\ \mathrm{svec}(\Delta Z) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}. \qquad (3.1)$$

On note que le système (3.1) est un système linéaire dont la matrice bloque est une matrice carrée d'ordre $(m + 2\tilde{n})$.

Remarque 3.1.1 *Afin d'obtenir la direction* $(\Delta X; \Delta y; \Delta Z)$; *on résoud le système précédent et on obtient* $(\mathrm{svec}(\Delta X); \Delta y; \mathrm{svec}(\Delta Z))$. *Puis, on introduit* smat *l'application réciproque de* svec *à ce dernier résultat.*

3.2 L'existence et l'unicité de la direction du déplacement

3.2.1 Condition suffisante

Pour étudier l'existence et l'unicité de la solution du système (3.1), on a besoin des lemmes suivants :

Lemme 3.2.1 *Soit* Q *un opérateur linéaire de* S^n *dans* S^n; *auto-adjoint et monotone, alors la matrice* $Q \in S^{\tilde{n}}_+$.

Preuve : Soit Q un opérateur linéaire auto-adjoint et monotone, alors

$$X \bullet Q(X) \geq 0; \forall X \in S^n.$$

D'après les propriétés de svec, on a :

$$\begin{aligned} X \bullet Q(X) &= \mathrm{svec}^T(X)\,\mathrm{svec}(Q(X)) \geq 0; \forall X \in S^n. \\ &= \mathrm{svec}^T(X)\,Q\,\mathrm{svec}(X) \geq 0; \forall X \in S^n. \end{aligned}$$

Par conséquent, la matrice $Q \in S^{\bar{n}}_{+}$. Ce qui achève la preuve. ■
Le prochain lemme montre que les deux matrices E et F sont inversibles.

Lemme 3.2.2 *Soit* $Z \in S^{n}_{++}$, *alors* $E \in S^{\bar{n}}_{++}$ *et* F *est inversible.*

Preuve : Soient $Z \in S^{n}_{++}$, $E = (Z \otimes_s I_n)$, $F = (ZP \otimes_s P)$. Puisque $Z, I_n \in S^{n}_{++}$ alors $E \in S^{\bar{n}}_{++}$ (d'après les propriétés de produit de Kronecker symétrisé). D'autre part on a :

$$\begin{aligned} F &= (ZP \otimes_s P), \\ &= (Z \otimes_s I_n)(P \otimes_s P), \\ &= E(P \otimes_s P). \end{aligned}$$

Pour montrer que F est inversible, on montre que si $Fx = 0$, alors $x = 0$. On pose

$$(P \otimes_s P)x = r.$$

Alors

$$Fx = (Z \otimes_s I_n)r = Er = 0 \Rightarrow r = 0 \text{ car } E \in S^{\bar{n}}_{++},$$

c-à-d : $(P \otimes_s P)x = 0$ et comme $P \in S^{n}_{++}$, alors

$$(P \otimes_s P) \in S^{\bar{n}}_{++}, \tag{3.2}$$

donc elle est inversible et

$$(P \otimes_s P)x = 0 \Rightarrow x = 0.$$

Par conséquent F est inversible. Ce qui achève la preuve. ■

Lemme 3.2.3 *Soient* $Q \in S^{\bar{n}}_{+}$, P *et* $Z \in S^{n}_{++}$. *Alors* $(E + FQ)$ *est inversible*.

Preuve : Soient $P, Z \in S^{n}_{++}$. On a :

$$\begin{aligned} (E + FQ) &= (Z \otimes_s I_n) + (ZP \otimes_s P)Q, \\ &= (Z \otimes_s I_n) + (Z \otimes_s I_n)(P \otimes_s P)Q, \\ &= (Z \otimes_s I_n)[I_{\bar{n}} + (P \otimes_s P)Q]. \end{aligned}$$

Pour montrer que $(E + FQ)$ est inversible, on montre que $(E + FQ)x = 0 \Rightarrow x = 0$. On pose

$$[I_{\tilde{n}} + (P \otimes_s P)Q]x = r.$$

Alors

$$(E + FQ)x = (Z \otimes_s I_n)r = Er = 0 \Rightarrow r = 0, \text{ car } E \in S^{\tilde{n}}_{++},$$

c-à-d : $[I_{\tilde{n}} + (P \otimes_s P)Q]x = 0$. On pose $(P \otimes_s P) = K$, puisque $P \in S^n_{++}$, alors d'après (3.2) $K \in S^{\tilde{n}}_{++}$ et comme $Q \in S^{\tilde{n}}_{+}$, la marice $[I_{\tilde{n}} + (P \otimes_s P)Q] = [I_{\tilde{n}} + KQ]$ est définie positive car :

$$KQ \sim K^{\frac{1}{2}}QK^{\frac{1}{2}} \text{ et } Q \in S^{\tilde{n}}_{+} \text{ donc } KQ \text{ est semi définie positive.}$$

Alors $[I_{\tilde{n}} + (P \otimes_s P)Q]$ est inversible et par conséquent

$$[I_{\tilde{n}} + (P \otimes_s P)Q]x = 0 \Rightarrow x = 0,$$

donc $(E + FQ)$ est inversible. Ce qui achève la preuve. ■

Maintenant, on va présenter une condition suffisante pour assurer l'existence et l'unicité de la direction de Nesterov-Todd.

Théorème 3.2.1 *Le système linéaire (3.1) admet une solution unique* $(\Delta y, \Delta X, \Delta Z) \in \mathbb{R}^m \times S^n \times S^n$ *si la matrice* $(E + FQ)^{-1}F$ *est définie positive.*

Preuve : On considère le système homogène suivant :

$$\begin{bmatrix} 0 & A & 0 \\ A^T & -Q & I_{\tilde{n}} \\ 0 & E & F \end{bmatrix} \begin{bmatrix} \Delta y \\ \text{svec}(\Delta X) \\ \text{svec}(\Delta Z) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{3.3}$$

et on montre que sa solution unique est le vecteur nul. D'après l'équation (2), on a :

$$\text{svec}(\Delta Z) = Q\,\text{svec}(\Delta X) - A^T \Delta y, \tag{3.4}$$

et la dernière équation devient :

$$(E + FQ)\,\text{svec}(\Delta X) = FA^T \Delta y,$$

d'après le Lemme [3.2.3], la matrice $(E + FQ)$ est inversible, donc

$$\mathrm{svec}(\Delta X) = (E + FQ)^{-1} F A^T \Delta y; \tag{3.5}$$

alors, la première équation devient :

$$[A(E + FQ)^{-1} F A^T] \Delta y = 0; \tag{3.6}$$

donc il suffit de montrer que la matrice $\left[A(E + FQ)^{-1} F A^T\right]$ est définie positive. Pour cela, on considère $v \in \Re^m - \{0\}$; comme les matrices A_i sont linéairement indépendantes, alors A est de plein rang et par conséquent $A^T v \neq 0$. De plus, on sait que $(E + FQ)^{-1} F$ est une matrice définie positive, donc il est clair que :

$$v^T A(E + FQ)^{-1} F A^T v > 0; \forall v \in \Re^m - \{0\}; \tag{3.7}$$

c-à-d : $\left[A(E + FQ)^{-1} F A^T\right]$ est définie positive, donc elle est inversible. Par conséquent :

$$\Delta y = 0;$$

est la seule solution de (3.3) et cela implique que

$$\mathrm{svec}(\Delta X) = 0;$$

et ensuite

$$\mathrm{svec}(\Delta Z) = 0; \text{ d'après l'expression (3.4).}$$

Cela montre que le système (3.1) admet une solution unique, ce qui achève la preuve. ■

Remarque 3.2.1 *Le résultat de ce théorème généralise la condition suffisante obtenue par Todd, Toh et Tütüncü [24] en (SDP) pour assurer l'existence et l'unicité de la direction.*

Le corollaire suivant montre une condition suffisante pour que la matrice $(E + FQ)^{-1} F$ soit définie positive.

Corollaire 3.2.1 *Si* $Q \in S^{\tilde{n}}_+$, *alors la matrice* $(E + FQ)^{-1} F$ *est définie positive.*

Preuve : Soit $g \in \Re^{\tilde{n}} - \{0\}$; on doit montrer que $g^T (E + FQ)^{-1} F g > 0$, pour cela on considère $k = (E + FQ)^{-T} g$; d'après le Lemme [3.2.3] la matrice $(E + FQ)$ est inversible donc $\det(E + FQ) \neq 0$ et $g \neq 0$ alors $k \neq 0$ et on a :

$$
\begin{aligned}
g^\top (E + FQ)^{-1} F g &= k^\top F (E + FQ)^\top k; \\
&= k^\top F (E + QF^\top) k; \\
&= k^\top (ZP \otimes_s P)[(Z \otimes_s I_n) + Q(PZ \otimes_s P)]k; \\
&= k^\top (ZP \otimes_s P)(Z \otimes_s I_n)k + k^\top (ZP \otimes_s P)Q(PZ \otimes_s P)k; \\
&= k^\top (Z \otimes_s I_n)(P \otimes_s P)(I_n \otimes_s Z)k + k^\top FQF^\top k; \\
&= k^\top E(P \otimes_s P)E^\top k + k^\top FQF^\top k; \\
&> k^\top FQF^\top k; \\
&\geq 0;
\end{aligned}
$$

la stricte inégalité est satisfaite car $(P \otimes_s P) \in S^{\hat{n}}_{++}$. En effet, $k \neq 0$ et d'après le Lemme [3.2.2], la matrice $E \in S^{\hat{n}}_{++}$ donc $k^\top E \neq 0$ et puisque $P \in S^n_{++}$ donc $(P \otimes_s P) \in S^{\hat{n}}_{++}$. La dernière inégalité est satisfaite car $k \neq 0$ et d'après le Lemme [3.2.2], la matrice F est inversible donc $k^\top F \neq 0$ et $Q \in S^{\hat{n}}_{+}$. Ce qui achève la preuve. ■

3.3 Résolution du système

Dans cette partie, on s'intéresse à la résolution du système (3.1) :

$$
\begin{cases}
A \,\mathrm{svec}(\Delta X) & = 0; \quad \cdots \ (1) \\
A^\top \Delta y - Q \,\mathrm{svec}(\Delta X) + I_{\hat{n}} \,\mathrm{svec}(\Delta Z) & = 0; \quad \cdots (2) \\
E \,\mathrm{svec}(\Delta X) + F \,\mathrm{svec}(\Delta Z) & = w; \quad \cdots (3)
\end{cases}
$$

d'après l'équation (2), on obtient :

$$\mathrm{svec}(\Delta Z) = Q \,\mathrm{svec}(\Delta X) - A^\top \Delta y; \tag{3.8}$$

en remplacant cette dernière dans l'équation (3), on obtient :

$$E \,\mathrm{svec}(\Delta X) + F \,[Q \,\mathrm{svec}(\Delta X) - A^\top \Delta y] = w;$$

qui est équivalente à :

$$(E + FQ) \,\mathrm{svec}(\Delta X) = w + FA^\top \Delta y; \tag{3.9}$$

d'après le Lemme [3.2.3], la matrice $(E+FQ)$ est inversible et par conséquent la solution $\mathrm{svec}(\Delta X)$ est donnée par :

$$\mathrm{svec}(\Delta X) = (E+FQ)^{-1}\left[FA^{\top}\Delta y + w\right]; \tag{3.10}$$

d'après (1); on a :

$$\begin{aligned} A\,\mathrm{svec}(\Delta X) = 0 \quad &\Leftrightarrow \quad A\left[(E+FQ)^{-1}\left[FA^{\top}\Delta y + w\right]\right] = 0;\\ &\Leftrightarrow \quad \left[A(E+FQ)^{-1}FA^{\top}\right]\Delta y + A(E+FQ)^{-1}w = 0;\\ &\Leftrightarrow \quad \left[A(E+FQ)^{-1}FA^{\top}\right]\Delta y = -A(E+FQ)^{-1}w; \end{aligned} \tag{3.11}$$

on pose :

$$(E+FQ)^{-1}w = u; \text{ où } u \in \mathbb{R}^{\tilde{n}};$$

alors

$$(E+FQ)u = w; \tag{3.12}$$

comme $(E+FQ)$ est inversible, on résoud ce système pour obtenir u.
Par conséquent, (3.11) devient :

$$\left[A(E+FQ)^{-1}FA^{\top}\right]\Delta y = -Au; \tag{3.13}$$

d'après (3.7) la matrice $\left[A(E+FQ)^{-1}FA^{\top}\right]$ est inversible, alors on résoud le système linéaire (3.13) et on obtient :

$$\Delta y = -\left[A(E+FQ)^{-1}FA^{\top}\right]^{-1}Au;$$

où

$$u = (E+FQ)^{-1}w;$$

puis on remplace ce résultat en (3.10) pour obtenir $\mathrm{svec}(\Delta X)$. De (3.8), on déduit que :

$$\mathrm{svec}(\Delta Z) = Q\,\mathrm{svec}(\Delta X) - A^{\top}\Delta y;$$

donc la solution du système (3.1) est donnée par :

$$\begin{cases} \Delta y = -\left[A(E+FQ)^{-1}FA^T\right]^{-1} A(E+FQ)^{-1}w; \\ \mathrm{svec}(\Delta X) = (E+FQ)^{-1}\left[FA^T\Delta y + w\right]; \\ \mathrm{svec}(\Delta Z) = Q\,\mathrm{svec}(\Delta X) - A^T\Delta y. \end{cases}$$

Remarque 3.3.1 *Pour définir la matrice* $\left[A(E+FQ)^{-1}A^T\right]$ *; on a besoin de calculer* $(E+FQ)^{-1}$*; et comme* $(E+FQ)$ *est inversible, donc elle est régulière et son inverse est calculé par la méthode de Gauss-Jordan.*

3.4 Calcul de la matrice P

Dans cette partie, on utilise les techniques introduite par Todd, Toh et Tütüncü pour calculer la matrice P (voir [18], [24]). Pour cela on a besoin du lemme suivant :

Lemme 3.4.1 *: Soit* $B \in S^n_{++}$ *telle que* $B = CC^T$*. Alors, la matrice* $W = C^{-1}B^{\frac{1}{2}}$ *est orthogonale.*

Preuve : On a $WW^T = C^{-1}BC^{-T} = C^{-1}CC^TC^{-T} = I$. Ce qui achève la preuve. ■

Maintenant, on considère la factorisation de Cholesky des deux matrices X et Z :

$$X = LL^T; \; Z = RR^T;$$

où L (respectivement, R) sont des matrices triangulaires inférieures et on considère maintenant la matrice R^TL où sa décomposition (SVD) est donnée par :

$$R^TL = U\Sigma V^T; \text{ t.q } \Sigma = \mathrm{diag}(\sigma_i) \text{ et } \sigma_i \text{ sont les valeurs singulières non nulles de } R^TL.$$

Posons $W = L^{-1}X^{\frac{1}{2}}$; d'après le Lemme [3.4.1] W est une matrice orthogonale, alors on obtient :

$$\begin{aligned} X^{\frac{1}{2}}ZX^{\frac{1}{2}} &= X^{\frac{1}{2}}L^{-T}(L^TRR^TL)L^{-1}X^{\frac{1}{2}}; \\ &= W^T(L^TRR^TL)W; \\ &= W^T(V\Sigma U^T)(U\Sigma V^T)W; \\ &= (W^TV)\Sigma^2(V^TW); \end{aligned}$$

telle que $\Lambda\Sigma^2\Lambda^T$ est la décomposition en valeurs propres de L^TZL. Comme $W^T\Lambda$ est orthogonale, on obtient :

$$(X^{\frac{1}{2}}ZX^{\frac{1}{2}})^{-\frac{1}{2}} = (W^T\Lambda)\Sigma^{-1}(\Lambda^TW);$$

alors

$$\begin{aligned}
P &= X^{\frac{1}{2}}(X^{\frac{1}{2}}ZX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}};\\
&= X^{\frac{1}{2}}(W^T\Lambda)\Sigma^{-1}(\Lambda^TW)X^{\frac{1}{2}};\\
&= X^{\frac{1}{2}}X^{\frac{1}{2}}L^{-T}(\Lambda\Sigma^{-1}\Lambda^T)L^{-1}X^{\frac{1}{2}}X^{\frac{1}{2}};\\
&= XL^{-T}(\Lambda\Sigma^{-1}\Lambda^T)L^{-1}X;\\
&= LL^TL^{-T}(\Lambda\Sigma^{-1}\Lambda^T)L^{-1}LL^T;\\
&= L\Lambda\Sigma^{-1}\Lambda^TL^T = GG^T.
\end{aligned}$$

Où $G = L\Lambda\Sigma^{-\frac{1}{2}}$.

3.5 Calcul de la matice V

D'après (2.6) on a :

$$V = \frac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}}DZD;$$

Afin de calculer la matrice V, on a besoin de calculer D telle que $D = P^{\frac{1}{2}}$. Puisque $P \in S^n_{++}$, alors P est diagonalisable et ses valeurs propres sont réelles positives c-à-d, il existe une matrice unitaire U telle que

$$P = U\mathrm{diag}(\lambda_1, \dots, \lambda_n)U^T; \text{ avec } \lambda_i \in sp(P);$$

et

$$D = P^{\frac{1}{2}} = U\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})U^T; \text{ donc } P^{\frac{1}{2}} \in S^n_{++}.$$

Par conséquent

$$V = \frac{1}{\sqrt{\mu}}DZD.$$

Chapitre 4

Tests numériques

Dans ce chapitre, on s'intéresse aux expériences numériques issues de l'application de l'Algorithme 2.4.4 sur des problèmes quadratiques semi-définis. On note par :

- $(X^0; y^0; Z^0)$: un point initial strictement réalisable et vérifie $\delta(X^0; Z^0; \mu^0) \leq \frac{1}{\sqrt{2}}$ avec $\mu^0 = \frac{X^0 \bullet Z^0}{n}$.
- $(X^*; y^*; Z^*)$: désigne la solution optimale du problème primal $(PQSDP)$ et dual $(DQSDP)$, respectivement.
- $iter$: désigne le nombre des itérations produites par l'Algorithme 2.4.4.
- m_P et m_D : désignent les valeurs optimales primales et duales des deux problème $(PQSDP)$ et $(DQSDP)$, respectivement.

4.1 Exemples

Exemple 1 *Soient* $n = 2m$; $Q(X) = 0_{n \times n}$; $C = -I_n$ *et* $\varepsilon = 10^{-4}$

$$A_i[j; k] = \begin{cases} 1 & si \;\; j = k = i \quad ou \;\; j = k = i + m\,; \\ 0 & \quad\quad\quad ailleurs, \end{cases}$$

et

$$b[i] = 2; \quad i = 1; \ldots; m;$$

la solution initiale primale-duale et strictement réalisable est donnée par :

$$X^0[i; j] = \begin{cases} 1.2 & si \;\; i = j = 1; \ldots; m; \\ 0.8 & si \;\; i = j = m + 1; \ldots; n; \\ 0 & ailleurs, \end{cases}$$

$$
Z^0[i;j] = \begin{cases} 1 & si \quad i = j = 1, \dots, n; \\ 0 & ailleurs, \end{cases}
$$

$$
y^0[i] = -2; \quad i = 1, \dots, m;
$$

avec

$$
\delta(X^0; Z^0; \mu^0) = 0.07335444 < \frac{1}{\sqrt{2}};
$$

la solution optimale primale-duale est donnée par :

$$
X^*[i;j] = \begin{cases} 1 & si \quad i = j = 1, \dots, n; \\ 0 & ailleurs, \end{cases}
$$

$$
Z^* = 0_{n \times n} \;\; et \;\; y^*[i] = -1; \quad i = 1, \dots, m.
$$

Pour $m = 2$ *et* $n = 4$; *notre programme quadratique semi-défini est donné par :*

$$
(PQSDP) \quad \begin{cases} \min[-x_{11} - x_{22} - x_{33} - x_{44}] \\ x_{11} + x_{33} = 2; \; x_{22} + x_{44} = 2; \\ X \in S^n_+; \end{cases}
$$

et son dual est :

$$
(DQSDP) \quad \begin{cases} \max[2y_1 + 2y_2] \\ y_1 + z_{11} = -1; \; y_1 + z_{33} = -1; \\ y_2 + z_{22} = -1; \; y_2 + z_{44} = -1; \\ Z \in S^n_+; \; y \in \mathbb{R}^2. \end{cases}
$$

On note que :

- la solution initiale est :

$$X^0 = \begin{bmatrix} 1.2 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.8 \end{bmatrix};\ Z^0 = I_4 \quad \text{et} \quad y^0 = [-2;\ -2]^\top.$$

- L'opérateur linéaire $Q(X) = 0_{4\times 4}$ est monotone et auto-adjoint, et par conséquent

$$\mathcal{Q} = 0_{10\times 10}.$$

- Les deux matrices A_1 et A_2 sont linéairement indépendantes telles que :

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- La solution optimal est :

$$X^* = I_4;\ Z^* = 0_{4\times 4} \ \text{ et } \ y^* = [-1;\ -1]^\top.$$

Le tableau suivant présente le nombre d'itérations, les valeurs optimales primales et duales et le saut de dualité pour différentes tailles du problème :

taille $(m;n)$	m_P	m_D	le saut de dualité	iter	temps (S)
(1, 2)	-2.00000000	-2.00006150	0.00006150	24	0.42
(2, 4)	-4.00000000	-4.00002503	0.00002503	37	0.46
(3, 6)	-6.00000000	-6.00001337	0.00001337	46	0.98
(4, 8)	-8.00000000	-8.00009938	0.00009938	53	6.31
(5, 10)	-10.00000000	-10.00005606	0.00005606	60	32.84
(6, 12)	-12.00000000	-12.00001260	0.00001260	67	162.12
(7, 14)	-14.00000000	-14.00009316	0.00009316	72	628.14

Remarque 4.1.1 *Dans tous les exemples suivants, les matrices* A_i *sont linéairement indépendantes et l'opérateur* $Q : S^n \to S^n$ *est auto-adjoint et monotone.*

Exemple 2 *Soient* $n = 2m;\ Q(X) = X\ ;\ C = -I_n$ *et* $\varepsilon = 10^{-4}$

$$A_i[j;k] = \begin{cases} 1 & \textit{si } j = k = i \quad \textit{ou } j = k = i + m\ ; \\ 0 & \textit{ailleurs,} \end{cases}$$

et

$$b[i] = 2; \quad i = 1;\dots;m;$$

la solution initiale est :

$$X^0[i;j] = Z^0[i;j] = \begin{cases} 1.2 & \textit{si } i = j = 1;\dots;m; \\ 0.8 & \textit{si } i = j = m+1;\dots;n; \\ 0 & \textit{ailleurs,} \end{cases}$$

$$y^0[i] = -1; \quad i = 1;\dots;m\,;$$

avec

$$\delta(X^0;Z^0;\mu^0) = 0.15655613 < \frac{1}{\sqrt{2}};$$

la solution optimale est :

$$X^*[i;j] = \begin{cases} 1 & \textit{si } i = j = 1;\dots;n; \\ 0 & \textit{ailleurs,} \end{cases}$$

$$Z^* = 0_{n\times n} \ \textit{ et } \ y^*[i] = 0; \quad i = 1;\dots;m.$$

Pour $m = 2$ *et* $n = 4$; *le programme quadratique semi-défini est donné par :*

$$(PQSDP) \begin{cases} \min[-x_{11} - x_{22} - x_{33} - x_{44} + \frac{1}{2}(x_{11}^2 + x_{22}^2 + x_{33}^2 + x_{44}^2)] \\ x_{11} + x_{33} = 2; \ x_{22} + x_{44} = 2; \\ X \in S_+^n; \end{cases}$$

et son dual est donné par :

$$
(DQSDP)\quad \begin{cases} \max[2y_1 + 2y_2 - \frac{1}{2}(x_{11}^2 + x_{22}^2 + x_{33}^2 + x_{44}^2)] \\ y_1 + z_{11} - x_{11} = -1; \quad y_1 + z_{33} - x_{33} = -1; \\ y_2 + z_{22} - x_{22} = -1; \quad y_2 + z_{44} + x_{44} = -1; \\ Z \in S^n_+;\ y \in \mathbb{R}^2. \end{cases}
$$

On note que :

- la solution initiale est :

$$
X^0 = Z^0 = \begin{bmatrix} 1.2 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.8 \end{bmatrix} \quad \text{et} \quad y^0 = [-1;\ -1]^T.
$$

- L'opérateur linéaire $Q(X) = X$ est monotone et auto-adjoint, et

$$
\tilde{Q} = (I_4 \otimes_s I_4).
$$

- Les matrices A_1 et A_2 sont :

$$
A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

- La solution optimale est :

$$
X^* = I_4;\ Z^* = 0_{4\times 4} \ \text{et}\ y^* = [0;\ 0]^T.
$$

taille $(m;n)$	m_P	m_D	le saut de dualité	iter	temps (S)
$(1,2)$	-1.0000000	-1.00009126	0.00009126	24	0.54
$(2,4)$	-2.0000000	-2.00009916	0.00009916	38	0.68
$(3,6)$	-3.0000000	-3.00008641	0.00008641	50	1.09
$(4,8)$	-4.0000000	-4.00008623	0.00008623	60	6.38
$(5,10)$	-5.0000000	-5.00008595	0.00008595	69	39.39

Exemple 3 *Soient* $m = 1;\ n = 2$, $\epsilon = 10^{-8};\ Q(X) = 0_{2\times 2}$ *et*

$$C = diag[3;\ -1];\ A_1 = diag[1;\ 2];\ b = [2];$$

la solution initiale est :

$$X^0 = diag[0.3000;\ 0.8500];$$

$$Z^0 = diag[4.0000;\ 1.0000];$$

$$y^0 = [-1.0000];$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.06199349 < \frac{1}{\sqrt{2}};$$

et

$$Q = 0_{3\times 3}.$$

Les résultats numérique

$$X^* = diag[0.0000;\ 1.0000];$$

$$Z^* = diag[3.5000;\ 0.0000];$$

$$y^* = [-0.5000].$$

Le tableau suivant présente le nombre d'itérations, les valeurs optimales primales et duales et le saut de dualité :

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
-0.9999999952	*-1.0000000047*	*0.0000000094*	*45*	*0.71*

Exemple 4 *Soient* $m = 2$; $n = 3$, $\epsilon = 10^{-8}$; $Q(X) = HXH$ *telle que :*

$$H = diag\left[\sqrt{2}; \sqrt{2}; 0\right];$$

et

$$C = diag[-2; -4; 0]; \qquad b = [1; 2]^{\top};$$

$$A_1 = diag[-1; 1; 0]; \quad A_2 = diag[1; 1; 1];$$

la solution initiale est :

$$X^0 = diag[0.4000; 1.4000; 0.2000];$$

$$Z^0 = diag[0.3000; 0.3000; 1.5000];$$

$$y^0 = [0.0000; -1.5000]^{\top};$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.2797794 < \frac{1}{\sqrt{2}};$$

et

$$Q = (H \otimes_s H).$$

Les résultats numériques

$$X^* = diag[0.5000; 1.5000; 0.0000];$$

$$Z^{*} = diag\,[0.0000;\ 0.0000;\ 1.0000]\,;$$

$$y^{*} = [0.0000;\ -1.0000]^{\top}\,.$$

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
-4.4999999984	*-4.5000000046*	*0.0000000062*	*55*	*0.99*

Exemple 5 *Soient* $m = 2;\ n = 3,\ \epsilon = 10^{-5};\ Q(X) = HXH$ *telle que :*

$$H = diag\left[\sqrt{2};\ \sqrt{2};\ 0\right]\,;$$

et

$$C = diag\,[0;\ 0;\ 0]\,; \qquad b = [4;\ 8]^{\top}\,;$$

$$A_1 = diag\,[1;\ 1;\ 0]\,; \quad A_2 = diag\,[1;\ 2;\ 1]\,;$$

la solution initiale est :

$$X^0 = diag\,[2.5000;\ 1.5000;\ 2.5000]\,;$$

$$Z^0 = diag\,[3.0000;\ 2.0000;\ 1.0000]\,;$$

$$y^0 = [3.0000;\ -1.0000]^{\top}\,;$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.22276582 < \frac{1}{\sqrt{2}};$$

et

$$Q = (H \otimes_s H).$$

Les résulats numériques

$$X^* = diag[2.0000;\ 2.0000;\ 2.0000]\,;$$

$$Z^* = diag[0.0000;\ 0.0000;\ 0.0000]\,;$$

$$y^* = [4.0000;\ 0.0000]^\top.$$

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
7.99999999	*7.99999177*	*0.00000822*	*40*	*0.99*

Exemple 6 *Soient* $m = 2;\ n = 3,\ \varepsilon = 10^{-6};\ Q(X) = HXH$ *telle que :*

$$H = diag\left[\sqrt{2};\ \sqrt{2};\ 0\right];$$

et

$$C = diag[0;\ 0;\ 2]\,; \qquad b = [6;\ 4]^\top;$$

$$A_1 = diag[1;\ 1;\ 1]\,; \quad A_2 = diag[-1;\ 1;\ 1]\,;$$

la solution initiale est donnée par :

$$X^0 = diag[1.0000;\ 1.5000;\ 3.5000]\,;$$

$$Z^0 = diag[1.0000;\ 2.0000;\ 1.0000]\,;$$

$$y^0 = [1.0000;\ 0.0000]^\top;$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.3038465 < \frac{1}{\sqrt{2}};$$

et

$$Q = (H \otimes_s H).$$

Les résultats numériques

$$X^* = diag[1.0000;\ 1.0000;\ 4.0000],$$

$$Z^* = diag[0.0000;\ 0.0000;\ 0.0000],$$

$$y^* = [2.0000;\ 0.0000]^T.$$

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
9.99999999	*9.99999913*	*0.00000086*	*45*	*0.88*

Exemple 7 *Soient* $m = 2$; $n = 4$, $\epsilon = 10^{-4}$; $Q(X) = HXH$ *telle que :*

$$H = diag\left[\sqrt{2};\ \sqrt{2};\ \sqrt{2};\ \sqrt{2}\right],$$

et

$$C = diag[-3;\ -10;\ 0;\ 0], \qquad b = [2;\ 11]^T,$$

$$A_1 = diag[-1;\ 1;\ 1;\ 0], \quad A_2 = diag[2;\ 3;\ 0;\ 1],$$

la solution initiale est donnée par :

$$X^0 = diag[1.5000;\ 2.0000;\ 1.5000;\ 2.0000],$$

$$Z^0 = diag[2.0000;\ 2.0000;\ 5.0000;\ 6.0000],$$

$$y^0 = [-2.0000;\ -2.0000]^T,$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.31154355 < \frac{1}{\sqrt{2}};$$

et

$$Q = (H \otimes_s H).$$

Les résultats numériques

$$X^* = diag\,[1.0000;\ 3.0000;\ 0.0000;\ 0.0000]\,;$$

$$Z^* = diag\,[0.0000;\ 0.0000;\ 1.0001;\ 1.0000]\,;$$

$$y^* = [-1.0000;\ -1.0000]^\top .$$

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
-22.99996579	*-23.00003818*	*0.00007239*	*45*	*1.28*

Exemple 8 *Soient* $m = n = 2$, $\varepsilon = 10^{-9}$; $Q(X) = HXH$ *telle que :*

$$H = \begin{bmatrix} 0.0125 & 0 \\ 0 & 0.0001 \end{bmatrix};$$

et

$$C = \begin{bmatrix} -0.00015625 & 0.0000012498 \\ 0.0000012498 & -0.00000001 \end{bmatrix}; \quad b = [1;\ 1]^\top;$$

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix};$$

la solution initiale est donnée par :

$$X^0 = \begin{bmatrix} 1 & 0.222222 \\ 0.222222 & 1 \end{bmatrix}; \ Z^0 = \begin{bmatrix} 0.666667 & 0.0000015276 \\ 0.0000015276 & 0.222222 \end{bmatrix};$$

$$y^0 = [-0.666667;\ -0.222222]^\top;$$

avec

$$\delta(X^0; Z^0; \mu^0) = 0.29383939 < \frac{1}{\sqrt{2}};$$

et

$$Q = (H \otimes_s H).$$

Les résultats numériques

$$X^* = \begin{bmatrix} 1 & -0.9876 \\ -0.9876 & 1 \end{bmatrix};\quad Z^* = \begin{bmatrix} 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \end{bmatrix};$$

$$y^* = [0.0000;\ 0.0000]^\top.$$

m_P	m_D	*le saut de dualité*	*iter*	*temps (S)*
-0.00007937941	*-0.00007938020*	*0.00000000078*	*49*	*0.93*

Conclusion générale et perspectives

La méthode de trajectoire centrale de type primal-dual est é¢cace pour résoudre plusieurs problèmes d'optimisations. Notamment, la programmation quadratique semi-dé…nie (QSDP), car au niveau de la théorie, on obtient une meilleure complexité polynomiale qui est de l'ordre $O(\sqrt{n}\log(\frac{n\mu_0}{\epsilon}))$: Mais, l'absence de la symétrie des directions et l'obtention d'un point strictement réalisable (qui n'est pas évoqué dans notre étude) avec la mesure de proximité $\delta(X;Z;\mu) < \frac{1}{\sqrt{2}}$ justi…ent la di¢culté de cette méthode.

Dans ce mémoire, on développe un algorithme de trajectoire centrale de type primal-dual à petit pas pour résoudre un problème quadratique semi-dé…ni (QSDP).

Notons que notre étude à réaliser les contributions suivantes :

- Etablir une condition su¢sante pour assurer la stricte faisabilité du pas de Newton complète.

- Etablir l'in‡uence du pas de Newton complète en saut de dualité.

- Etablir le taux de la réduction de μ pour réster proche de la trjectoire centrale dans le cas (QSDP) où les directions ne sont pas orthogonales.

- Etablir l'existence et l'unicité de la direction de déplacement de (NT) par une condition su¢sante, qui devient une généralisation du résultat obtenu par Todd, Toh et Tütüncü en (SDP).

Notre prochain travail concernant la résolution d'un problème (QSDP) par la méthode de trajectoire centrale basée sur les fonctions noyaux et une nouvelle transformation algébrique équivalente de puissance ($\psi(t) = \sqrt{t}$) appliquée à l'équation de centralité non linéaire avec d'autre schéma de symétrisations.

Annexe

Dans cette partie, on représente la preuve des Lemmes qui ont été utilisés en Chapitre 2 comme des lemmes techniques introduits par De klerk en (SDP) [7].
Les deux premiers lemmes nous conduit à obtenir le gradient de la fonction $f_\mu(X)$, où

$$f_\mu(X) = C \bullet X + \frac{1}{2} X \bullet Q(X) - \mu \log\det X; \ \mu > 0.$$

Lemme 2.4.1 Soit $f : S^n_{++} \to \mathbb{R}$ une fonction donnée par $f(X) = \log\det X$, et

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \cdots & \frac{\partial f(X)}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X)}{\partial x_{n1}} & \cdots & \frac{\partial f(X)}{\partial x_{nn}} \end{bmatrix};$$

alors $\nabla f(X) = X^{-1}$.
Preuve : Soient $X \in S^n_{++}$ et $H \in S^n$ avec $(X + H) \in S^n_{++}$. On a :

$$\begin{aligned} f(X + H) - f(X) &= \log\det(X + H) - \log\det X; \\ &= \log\left(\frac{\det(X + H)}{\det(X)}\right); \\ &= \log\left[\det(X^{-1})\det(X + H)\right]; \\ &= \log\left[\det((X^{-1})(X + H))\right]; \\ &= \log\det(I + X^{-1}H); \\ &= \log\det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}); \text{ car } X^{-1}H \sim X^{-\frac{1}{2}}HX^{-\frac{1}{2}}. \end{aligned}$$

Comme $X \in S^n_{++}$ et $H \in S^n$ telle que $(X + H) \in S^n_{++}$ alors :

$$\left[\prod_{i=1}^{n} \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}})\right]^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}); \forall \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) > 0;$$

vous pouvez consulter [13] pour voir la démonstration de cette inégalité, qui est équivalente à :

$$\det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \leq \left(\frac{1}{n}\mathrm{Tr}\,(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}})\right)^n.$$

En prenant le logarithme pour cette dernière, on obtient :

$$\begin{aligned}
\log\det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) &\leq \log\left(\frac{1}{n}\operatorname{Tr}(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}})\right)^n;\\
&= n\log\left(\frac{1}{n}\operatorname{Tr}(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}})\right);\\
&= n\log\left(1 + \frac{1}{n}\operatorname{Tr}(X^{-\frac{1}{2}}HX^{-\frac{1}{2}})\right).
\end{aligned}$$

On sait que $\log(1 + t) \leq t, \forall t > -1$, alors :

$$\begin{aligned}
f(X + H) - f(X) &\leq \operatorname{Tr}(X^{-\frac{1}{2}}HX^{-\frac{1}{2}});\\
&= \operatorname{Tr}(X^{-1}H);\\
&= X^{-1} \bullet H;
\end{aligned}$$

et par conséquent $\nabla f(X) = X^{-1}$, ce qui achève la preuve. ■

Lemme 2.4.2 Soit $f : S^n_{++} \to \mathbb{R}$ une fonction donnée par $f(X) = \operatorname{Tr}(CX) = C \bullet X$, où $C \in S^n$, alors $\nabla f(X) = C$.

Preuve : Soient $X \in S^n_{++}$ et $H \in S^n$ avec $(X + H) \in S^n_{++}$. On a :

$$\begin{aligned}
f(X + H) - f(X) &= \operatorname{Tr}(C(X + H)) - \operatorname{Tr}(CX);\\
&= \operatorname{Tr}(C(X + H) - CX);\\
&= \operatorname{Tr}(CX + CH - CX);\\
&= \operatorname{Tr}(CH) = C \bullet H;
\end{aligned}$$

et par conséquent $\nabla f(X) = C$. Ce qui achève la preuve. ■

Le résultat suivant conduit à obtenir le Hessien de la fonction $f_1(X)$.

Lemme 2.4.3 Soient $f : S^n_{++} \to \mathbb{R}$ une fonction donnée par $f(X) = \log\det X$ et $\nabla f(X) = X^{-1}$, alors $\nabla^2 f(X)$ est un opérateur linéaire qui vérifie

$$\nabla^2 f(X)H = -X^{-1}HX^{-1}, \forall H \in S^n.$$

Preuve : Soit $L(S^n, S^n)$, l'espace des opérateurs linéaires de S^n dans S^n. La dérivée de ∇f au sens de Frechet est définie par une fonction

$$\nabla^2 f : S^n \to L(S^n, S^n)$$

avec

$$\lim_{\|H\|\to 0} \frac{\left\| \nabla f(X+H) - \nabla f(X) - \nabla^2 f(X)H \right\|}{\|H\|} = 0. \tag{4.1}$$

On montre que $\nabla^2 f(X)H = -X^{-1}HX^{-1}$ vérifie (4.1).
Soit $H \in S^n$ telle que $(X+H)$ est inversible, on considère :

$$\begin{aligned}
\left\| \nabla f(X+H) - \nabla f(X) - \nabla^2 f(X)H \right\| &= \left\| (X+H)^{-1} - X^{-1} + X^{-1}HX^{-1} \right\|; \\
&= \left\| (X+H)^{-1}(I - (X+H)X^{-1} + (X+H)X^{-1}HX^{-1}) \right\| \\
&= \left\| (X+H)^{-1}(HX^{-1}HX^{-1}) \right\|; \\
&\le \left\| (X+H)^{-1} \right\| \left\| HX^{-1}HX^{-1} \right\|; \\
&\le \left\| (X+H)^{-1} \right\| \|H\| \left\| X^{-1}HX^{-1} \right\|;
\end{aligned}$$

c-à-d $\nabla^2 f(X)H$ vérifie (4.1) $\Rightarrow$ $\nabla^2 f(X)H = -X^{-1}HX^{-1}$. Ce qui achève la preuve. ■
les deux Lemmes suivants sont utilisés pour montrer la stricte faisabilité du pas de Newton complète.

Lemme 2.4.4 [7, lemme 3.3.1] Soient $X(\alpha) := X + \alpha\Delta X$; $Z(\alpha) := Z + \alpha\Delta Z$ telles que $X \succ 0$ et $Z \succ 0$. Si

$$\det(X(\alpha)Z(\alpha)) > 0; \forall 0 \le \alpha \le \bar{\alpha}.$$

Alors : $X(\bar{\alpha}) \succ 0$ et $Z(\bar{\alpha}) \succ 0$.

Preuve : Soient $X \succ 0$, $Z \succ 0$ et

$$\begin{aligned}
f : \ [0;\bar{\alpha}] \times \mathbb{R} &\to \mathbb{R} \\
(\alpha;\lambda) &\mapsto f(\alpha;\lambda) = \det(X(\alpha) - \lambda I);
\end{aligned}$$

où f est continûment différentiable $\forall \alpha \in]0;\bar{\alpha}[$ et $\lambda \in \mathbb{R}$. On considère $(\alpha;\lambda^*)$ telle que $\lambda^* \in sp(X(\alpha))$; alors

$$f(\alpha;\lambda^*) = 0;$$

d'après le Théorème des fonctions implicites, il existe deux voisinages $U; V$ de α et λ^*; respectivement et une application φ de U dans V; $\varphi \in C^k$; telle que pour tout $(\alpha;\lambda) \in U \times V$, on a :

$$f(\alpha;\lambda) = 0 \Leftrightarrow \lambda = \varphi(\alpha);$$

avec $\lambda \in sp(X(\alpha))$; où la fonction $\varphi(\alpha)$ vérifie les propriétés suivantes :

- φ est définie et continue sur $[0;\bar{\alpha}]$.
- $\varphi(0) = \lambda > 0$, car par l'hypothèse on a $X(0) = X \succ 0$, donc $sp(X) \subset]0;+\infty[$.
- $\varphi(\alpha) \neq 0, \forall \alpha \in [0;\bar{\alpha}]$; car par l'hypothèse on a :

$$\det(X(\alpha)Z(\alpha)) = \det(X(\alpha))\det(Z(\alpha)) > 0, \forall \alpha \in [0;\bar{\alpha}].$$

Ce qui implique $0 \notin sp(X(\alpha))$ et $0 \notin sp(Z(\alpha)), \forall \alpha \in [0;\bar{\alpha}]$. Par conséquent,

$$\varphi(\alpha) > 0, \forall \alpha \in [0;\bar{\alpha}] ;$$

alors $sp(X(\alpha)) > 0, \forall \alpha \in [0;\bar{\alpha}]$; ce qui est équivalent à : $X(\alpha) \succ 0, \forall \alpha \in [0;\bar{\alpha}]$; avec la même démonstration pour la matrice $Z(\alpha)$; on obtient $X(\bar{\alpha}) \succ 0$ et $Z(\bar{\alpha}) \succ 0$. Ce qui achève la preuve. ■

Lemme 2.4.5 [7, lemme 3.3.3] Soient $Q \in S^n$ et $M \in \mathbb{R}^{n\times n}$ une matrice anti-symétrique. Si $Q \succ 0$ alors $\det(Q+M) > 0$. De plus, si on $\lambda_i(Q+M) \in \mathbb{R}$ et $i = 1, \ldots, n$, alors

$$0 < \lambda_{min}(Q) \leq \lambda_{min}(Q+M) \leq \lambda_{max}(Q+M) \leq \lambda_{max}(Q).$$

Preuve : Soient $Q \in S^n_{++}$ et $M \in \mathbb{R}^{n\times n}$ une matrice anti-symétrique. Pour tout $x \in \mathbb{R}^n - \{0\}$ on a :

$$x^\top(Q+M)x = x^\top Qx > 0, \ (\text{car } x^\top Mx = 0);$$

donc

$$\det(Q+M) \neq 0;$$

et par conséquent

$$\det(Q+tM) \neq 0, \forall t \in \mathbb{R}, \ (\text{car } tM \text{ est anti-symétrique}).$$

Maintenant, on pose :

$$\psi(t) = \det(Q+tM), \forall t \in \mathbb{R}.$$

La fonction $\psi(t)$ vérifie les propriétés suivantes :

- $\varphi(0) = \det(Q) > 0.$
- $\varphi(t)$ est continue par rapport à $t, \forall t \in \mathbb{R}$ car $\det(Q + tM)$ s'écrit sous la forme d'un polynôme par rapport à t.
- $\varphi(t) \neq 0, \forall t \in \mathbb{R}.$

Alors

$$\det(Q + tM) > 0, \forall t \in \mathbb{R},$$

et pour $t = 1$, on obtient $\det(Q + M) > 0.$
Pour montrer la deuxième partie du lemme, on suppose qu'il existe $\lambda > 0$ une valeur propre de $(Q + M)$, telle que $\lambda > \lambda_{max}(Q)$. Alors, $(\lambda I - Q) \in S^n_{++}$, ce qui implique que $[(Q + M) - \lambda I]$ est régulière c-à-d $\det[(Q + M) - \lambda I] \neq 0$ contradiction car $\lambda \in sp(Q + M)$; donc $(Q + M)$ n'admet pas une valeur propre (λ) supérieure à $\lambda_{max}(Q)$, donc

$$\lambda_{max}(Q + M) \leq \lambda_{max}(Q).$$

Par l'analogie, on suppose qu'il existe $\lambda > 0$ une valeur propre de $(Q + M)$ telle que $\lambda < \lambda_{min}(Q)$, donc $(Q - \lambda I) \in S^n_{++}$. avec le même argument, on a : $\det[(Q + M) - \lambda I] \neq 0$, c-à-d $\lambda \notin sp(Q + M)$ contradiction car $\lambda \in sp(Q + M)$; donc $(Q + M)$ n'admet pas une valeur propre (λ) inférieure à $\lambda_{min}(Q)$. Finalement, on obtient :

$$0 < \lambda_{min}(Q) \leq \lambda_{min}(Q + M) \leq \lambda_{max}(Q + M) \leq \lambda_{max}(Q).$$

Ce qui achève la preuve. ■

Bibliographie

[1] M. Achache. A new primal-dual path-following method for convex quadratic programming. Computational and Applied Mathematics.(25) pp. 97-110 (2006).

[2] M. Achache. A weighted path-following method for the linear complementarity problem. Universitatis Babes. Bolyai. Series Informatica (49) (1) pp. 61-73 (2004).

[3] M. Achache. Complexity analysis and numerical implementation of a short-step primal-dual algorithm for linear complementarity problems. Computational and Applied Mathematics. (216), 1889-1895, (2010).

[4] L. Adler and R.D.C. Monteiro. Interior path-following primal-dual algorithm . Part II: convex quadratic programming. Mathematical programming, (44), 43-66, (1989).

[5] F. Alizadeh. Interior point methods in semide…nite programming with application to combinatorial optimization. SIAMJ. Optim. (5), 13-51, (1995).

[6] C. Besse. Introduction à l'analyse matriciel. Laboratoire MIP. Université Paul Sabatier. Septembre (2003).

[7] E.De Klerk. Interior point methods for semide…nite programming. Master of Science in the Faculty of Enginnering University of Pretoria, (1997).

[8] E.De Klerk, C. Roos, T. Terlaky. A short survey on semide…nite programming. Faculty of Technical Mathematics and Informatics, Delft University of Technology. Netherlands (1997).

[9] E. De Klerk, C. Roos, T. Terlaky. On primal-dual path-following algorithms for semide…nite programming. Mathematical programming, 137-157, (1998).

[10] N.J. Higham. Computing the nearest correlation matrix ¡ a problem from …nance. IMA journal of numerical analysis 22, 329-343, (2002).

[11] R.A. Horn, R.J. Johnson. Matrix analysis, Compridge University Press, UK, (1986).

[12] B. Jansen, C. Roos, T. Terlaky and J. Ph. Vial, Primal-dual algorithm for linear programming based on the logarithmic barrier method. J. of optimization. Theory and applications, 83: 1-26, (1994).

[13] A. Keraghel. Analyse convexe. Theorie fandamentale et exercices. Université de Sétif. Faculté des sciences. Laboratoire de mathématique fondamentale et numérique. (2001).

[14] M. Kojima, M. Shida, S. Shindoh. Search directions in the SDP and monotone SDLCP: Generalization and inexact computation. Mathematical programming 85 pp. 51-80. (1999).

[15] B. Krislock. Numerical solution of semide…nite constrained least squares problems. University Regina, (2000).

[16] R. Leroy. Polynômes positifs, sommes de carrés Programmation semi-dé…nie positive Application à la minimisation. mémoire DEA, univérsité de Rennes. (2004).

[17] J.W. Nie, Y.X. Yuan. A potential reduction algorithm for an extended SDP problem (work supported by Chinese NNSF). Institute of Computational Mathematics and Scienti…c/Engineering Computing, (1999).

[18] J.W. Nie, Y.X. Yuan. A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps. Annals of Operations Research 103, 155-133, (2001).

[19] J. Peng, C. Roos, T. Terlaky. New complexity analysis of the primal-dual method for semide…nite optimization based on the Nesterov-Todd direction. Journal of optimization theory and applications; vol.109, no.2,pp. 327-343, (2001).

[20] W.H. Press, B.P. Flannery, S.A. Tenkolsky, W.T. Vetterling. Numerical Recipes in Pascal. The Art of Scienti…c computing. Combridge University Prees (1987).

[21] C. Roos, T. Terlaky and J. Ph. Vial, Theory and algorithms for linear optimization. An interior point approach. John-Wiley. Sons, Chichester, UK, (1997).

[22] J.F. Sturm and S. Zhang. Symmetric primal-dual path following algorithms for semide…nite programming. Technical Report 9554/A, Tinbergen Institute, Erasmus University Rotterdam, (1995).

[23] M.J. Todd. A study of search directions in primal-dual interion-point methods for semide…nite programming.Working paper, school of OR and Industrial Engineering, Cornell University, Ithaca. New-York 14853,(1998).

[24] M.J. Todd, K.C. Toh, R.H. Tütüncü. On the Nesterov-Todd direction in semide…nite programming. Mathematical programming. Copyright (c) by the Society for Industial and Applied Mathematics. SIAM Journal on Optimization 8: 769-796, (1998).

[25] K.C. Toh. An inexact primal-dual path following algorithm for convex quadratic SDP. Mathematical programming 112.pp. 221-254, (2008).

[26] K.C. Toh, R.H. Tütüncü, M.J. Todd. Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems. Paci…c journal of optimization. (3), $N^0$1, pp. 135-164, (2007).

[27] G.Q. Wang,Y.Q. Bai. primal-dual interior point algorithm for convex quadratic semide…nite optimization. Nonlinear analysis, (2009).

[28] G.Q. Wang and Y.Q. Bai. A new primal-dual path-following interior point algorithm for semide…nite programming. Journal of Mathematical Analysis and Applications, (353), 339-349, (2009).

[29] G. Wang, D. Zhu. A uni…ed Kernel function approach to primal-dual interior-point algorithms for convex quadratic SDO. Copyright (c) by Spriger Science+Business Media, LLC (2010).

[30] S.J. Wright. Primal-dual interior point methods. copyright by SIAM. (1997).

[31] Y. Ye. Interior point algorithms. Theory and Analysis. John-Wiley. Sons, Chichester, UK, (1997).